POPULATION DYNAMICS OF MULTIPLE VIRAL INFECTIONS

A Dissertation
Presented to
The Academic Faculty

By

Bradford Taylor

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Physics

Georgia Institute of Technology

December 2016

POPULATION DYNAMICS OF MULTIPLE VIRAL INFECTIONS

Approved by:

Dr. Joshua S. Weitz, Advisor
School of Biological Sciences and
School of Physics
Georgia Institute of Technology

Dr. Kurt Wiesenfeld
School of Physics
Georgia Institute of Technology

Dr. Flavio Fenton
School of Physics
Georgia Institute of Technology

Dr. Sam Brown
School of Biological Sciences
Georgia Institute of Technology

Dr. Matthew Sullivan
Department of Microbiology
Ohio State University

Date Approved: August 26, 2016

To me, absurdity is the only reality.

Frank Zappa

ACKNOWLEDGEMENTS

I would like to thank my advisor, Joshua Weitz, for his guidance and support throughout my PhD. His hard work and dedication helped transform a biologically-naive, liberal arts physics graduate with a penchant for overly-lofty sentences rife with parentheticals to the shrewd, concise communicator exemplified from here-on-out (what a success!). Josh's humanity made the overall PhD process far from as psychologically damaging as possible. He also inspires hope on the soccer field as my body proceeds through the inevitable decline of senescence. In all seriousness, Josh is an exceptional mentor – one who has not forgotten what it's like to be a grad student.

I would like to thank those that directly acted as mentors throughout my PhD experience. Tae Lee had the lucky experience of guiding me on my first project, which involved a primer on how to write articles–namely, through insistence, what does and does not belong in a figure caption. Mike Cortez provided welcome mathematically-rigorous feedback and guidance especially after my departure from focusing on pure mathematics in undergrad.

I would like to also thank my close collaborators. Catherine Penington has been gracious with our regular Skype meetings despite something like a 12 hour time difference. Jonathan Dushoff is an inspiration given his commitment, or natural disposition, towards humorous quips. Will Ratcliff and Jenn Pentz are doing the real work as experimentalists: I thank them for the chance to work with their awesome system, seriously if you don't know about snowflake yeast go look it up right now. Matthias Fischer is also doing the real work as an experimentalist, but he deserves particularly praise. Matthias is a modern day scientific pioneer in that he focuses on a non-model organism. I hope the scientific community as whole recognizes this value and supports his future work along with others who are exploring the full diversity of life.

I would like to thank all of the members of the Weitz lab past and present (and future if you cite my work!). More specifically, I would like to thank Abhi Das, Richard

Joh, Chuck Price (who was the first to make fun of me numerous times in the lab space), Kristen Knipe, Gabriel Mitchell (who inspired me to ask increasingly more ridiculous questions and who validated the fact that arguments over programming languages are largely arbitrary), Lauren Childs, Joey Leung, Ceyhun Eksin, Hayriye Gulbadek, Shengyun Peng, Keith Paarporn, Shimantika Sharma, Ashley Coenen, Yu-hui Lin, Devika Singh, Walker Gussler, and Nanditha Rajamani. Special thanks goes out to Alex Bucksh who's always up for a beer, watching soccer, and talking about point-set topology. The most thanks goes to Luis Jover–my PhD kindred spirit and the only person who could deal with my "humor" for 6 years.

All work and no play makes Jack a dull boy, and so I would like to thank all of my social support outside the lab. I would particularly like to highlight fellow Physics PhD students: Jeffrey Tithof–whose labspace has excellent chairs to procrastinate on, Matt Kinsey–who acted as my springboard for my cynicism, Jim Waters–who will always get a Simpsons reference and Martin Anquez–who is an excellent target for my joking. I want to especially thank Chris Marcotte for participating in multiple non-IRB approved human experiments involving subjecting ourselves to toxic levels of Panda Express. I would also like to thank all of my PhD friends in Biology: Lava Rishishwar for an endless supply of bagels, Miguel Rodriguez for clogging up the cluster all the time, Linh Chau for always brightening my day, Will Overholt for allowing me to talk at him more than with him than I should.

On a very serious note I would like to thank my thesis committee. I would like to thank Sam Brown for the perspective he provides as a lynchpin between experimental and quantitative biology. Similarly, Matt Sullivan grounds this work and will push my further pondering on viral dynamics in marine systems. I would like to thank Flavio Fenton for providing his sharp knowledge of dynamics and his sharp wit. Lastly, I thank Kurt Wiesenfeld for the lucid perspective he provides as the guru on top of the physics mountain.

I would like to also thank the funding sources that financed my graduate studies. I would like to acknowledge the Nerem Travel Fellowship, which funded my visit to Matthias Fis-

cher, which led to the inclusion of experimental data in this thesis. I also acknowledge my PI's grants: NSF Biological Oceanography, the James S. McDonnell foundation, and the Burroughs Wellcome Fund.

Finally I would like thank my family. I love them and they love me too, even when I don't call often enough. Outside of their unconditional support, I thank my parents for the direction they provided. Discussions with my father guided the importance in looking at problems from the ground up, helping me ask why and how. Discussions with my mother fostered a more holistic approach that helped me consider the implications. Most important for my progress in science, they instilled a modesty that keeps my beliefs and understanding flexible. I thank my sister and her husband for wine, bourbon, and much-needed support in some of the more ridiculous moments of my life. My nephews are too young to get thanks for doing anything, but they sure are cute!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

At the base of life – there are microbes. These microbes are infected by viruses. When multiple viruses infect the same host cell they are able to interact and, in turn, alter infection dynamics. These changes can scale-up to have major effects on the ecology and evolution of viruses and their hosts. The scaling up of dynamics suggests that in order to fully understand how viruses drive global processes such as the biogeochemical cycle we must understand what role multiple infections play. In this thesis, we study the causes and consequences of multiple viral infections at a population level using both theoretical and experimental approaches. First, we study how clustering of populations in space leads to an increased frequency and severity of multiple viral infections. This suggests multiple infections may occur more often than previously believed given frequently observed patchiness of population in marine environments. Next, to address the consequences of multiple infection, we focus on a recently discovered model system for viral coinfection–the virophage. By studying this virus of a virus, we show that it increases host abundances while reducing viral abundances. Thus the effects of viral interaction during coinfection can reduce the predatory pressure of the viruses on the population level. We also address the role that space plays in mediating virophage coexistence given different modes of coinfection. This suggests how coinfection likely occurs in virophage systems and, thus, points forward to understand the influence of natural selection on this three-player system. Finally, we present empirical results of infection dynamics for a virus that utilizes a unique means of viral replication–the virus factory. We show multiple infections lead to a reduction in burst size. Additionally, we hypothesize a unique growth dynamic of reduced latency time as a result of multiple infection. Altogether, this work highlights how virus-host dynamics both drive and are driven by multiple infections.

CHAPTER 1
INTRODUCTION

Seminal work by Bergh et al in 1989 showed that viruses exist in abundance in marine environments [1]. Abundances commonly exceeded 10^6 and even 10^7 per ml, a result that has been confirmed by follow-up studies. This finding revolutionized understanding of the global role of viruses and jumpstarted the field of viral ecology [2, 3, 4, 5, 6]. We now understand viruses impact biogeochemical cycles by infecting roughly 20% of marine bacteria daily [7, 8]. These global processes result from the culmination of many individual infection events. Hence, changes to infection dynamics scale up to drive global processes.

The life-history traits of hosts and viruses associated with infection are typically thought of in terms of viral "hijacking" of the host cell. However, such two player interactions can also involve additional viruses. That is, interactions between viruses are possible when infecting the same host cell because, during infection, phage gene products diffuse in a cell creating a pool of resources available to more than one virus [9]. The type of infection can be classified based on what viruses are present. In this thesis we utilize the term multiple infection to refer to the broadest case of more than one virus infecting the same host cell. In addition, we utilize the term coinfection to denote the case when multiple viruses of different strains are present in the same cell. Coinfection can occur between competing viruses. In this case, neither virus relies on the presence of the other for reproduction. Alternatively, coinfection can occur with one virus being a hyper-parasite–its reproduction depends on the presence of the other virus. In this thesis we focus on multiple infections of the same strain and on coinfection of hyper-parasitic viruses known as virophage.

A possible net result of multiple infections is altered infection dynamics. For example, multiple infections can alter the timing of lysis [10] and the number of viruses produced [11, 12]. Additionally, multiple infections can even influence the decision of lysis in favor

of the virus entering a quasi-dormant state with its genome integrated into the genome of its host [13]. A bacterial cell whose genome contains a quasi-dormant temperate phage genome is termed a lysogen. Presumably, the effects on these individual level dynamics would scale up to affect the ecology between populations.

Multiple infections also affect the evolutionary dynamics of viruses. In particular with analogy to polyploidy, multiple infection diversifies the possible evolutionary dynamics. For example, multiple infection also allows sharing of genetic material between viruses via recombination [14, 15]. Certain viruses can exchange modular components of their modular morphology during viral production leading to chimeric viruses [16]. These chimeric viruses also exhibit phenotypic mixing which can alter phage bacteria interaction networks of who infects whom [17]. Together these phenomenon demonstrate that multiple infections alter dynamics and thus apply evolutionary pressures upon which natural selection can act.

In an ecological sense, the pool of resources shared between infecting viruses forms a temporary intracellular niche in which individual strains may specialize. Indeed, as each infecting virus both contributes and competes from the pool of resources a tragedy of the commons scenario may ensue [18, 19]. Experimental work has shown that evolution in high multiplicity of infection environments leads to a prisoner's dilemma scenario between coinfecting viruses [20]. Hence, multiple infections provide a means for viruses to engage (or disengage) in social cooperation/evolution. For example, the ϕ6 RNA virus amplifies rates of multiple infection by altering the host surface receptors following infection [21]. It was hypothesized that the benefit of such apparent increased "sociality" was to increase rates of evolutionary diversification. However, this benefit did not outweigh the detriments of the prisoner's dilemma scenario. This tension provides context for the evolution of responses that limit multiple infections such as super-infection exclusion [22, 23].

Other consequences of multiple infections can be learned from turning to a microscopic form of hyperparasitism: the virophage–a virus that parasitizes a giant virus. The discovery

of giant viruses shocked the virology community in 2003 [24]. Giant viruses are operationally defined as a virus with a smallest dimension that exceeds 300nm. In addition to physical sizes comparable to bacteria, giant viruses feature comparable genome sizes (e.g., 2.5 Mbp [25]). A larger genome affords giant virus increased complexity. Observations of infections by specific giant viruses, the Nucleocytoplasmic Large DNA viruses (NCLDV), led to another major surprise in virology–cytoplasmic virus factories [26]. Cytoplasmic virus factories initiate from the infecting viral particles [27]. Whereas viruses typically completely rely on their host for viral reproduction, replication of viruses can initiate in the viral core as a result of viral encoded transcriptional machinery [28]. Ultimately, the viral genetic material remains in the viral factory instead of entering the host nucleus [27]. Eventually, viral progeny emerge on the exterior of the virus factory (see Figure 1.1) [26]. Together, the independent nature of these giant viruses has led some to argue that giant viruses constitute a fourth domain of life [29] (although see [30] for a counter-argument).

Whether life or not, giant viruses that utilize viral factories feature a unique replication method. The virus factories of giant viruses are also unique insofar as they provide an opportunity for cheating by another kind of (surprise) virus – the virophage [31]. Viruses are selfish entities that rely on exploiting their host replicative machinery in order to propagate. The virus factory, common to NCLDV, provides an opportune niche for exploitation. When coinfecting the virophage interacts with the virus factory ultimately leading to the production of virophage on the surface of the virus factory. Coupled with the production of virophage is a sizable decrease in the burst size of its viral partner, e.g., estimates have placed this reduction at 70% [31]. Part of this reduction in burst size stems from virophage directly interfering with virus production causing misfolding and abnormal growth of the viral capsid [31]. These effects differentiate virophage from other biological entities that rely on coinfection to propagate such as defective interfering particles [32] and satellite viruses [33]. Thus their reliance on coinfection and the stark impact on their viral hosts suggests that this is a valuable system to study the ecological and evolutionary impact of

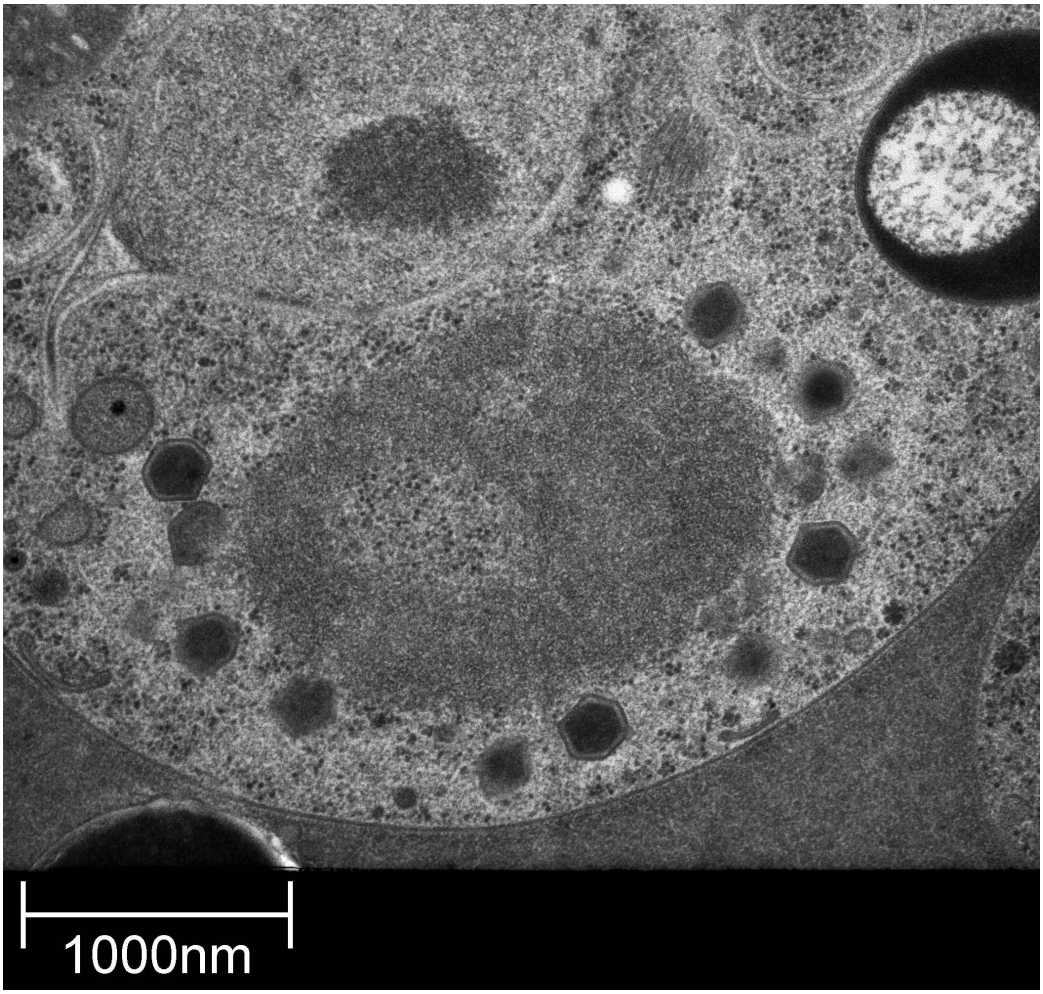


Figure 1.1: CroV viral factory at a later stage of infection of *Cafeteria roenbergensis*. The large grey mass at the center is the viral factory. The dark hexagons at the periphery of the viral factory are progeny CroV particles. Image credit: Matthias Fischer (MPI Heidelberg).

multiple infections.

The central theme of this thesis is on the role of multiple infections on the population dynamics of hosts and viruses. A central theme of the second and fourth chapters is how space affects host dynamics. The exploration of spatial dynamics in biology was initiated by Turing who argued that concentration gradients observed in a developing embryo could form from a separation of diffusion constants between chemical species [34]. Later work applied continuous, spatial models on ecological problems such as the emergence of planktonic patchiness [35]. These works, among other, firmly placed spatial ecology analysis in

the realm of pattern formation. However, with more powerful computers came the advent of individual-based simulations, in which there are effects of being discrete and spatial [36]. It became clear that not only could individual interactions lead to spatial patterning, but that clustering of populations could lead to novel impacts on interactions. It's with this motivation that we address spatial dynamics in this thesis: the coupling between spatial patterning and multiple infections.

The dynamic role of hyperparasitism forms a central theme of Chapters 3-5. Virophage coexist as hyper-parasites in a tri-partite system with a virus and a mutual host. Hyper-parasitism is prominent among insect communities [37]. The ecological importance of virophage remains uncertain. Metagenomic studies suggest that virophage are common in inland water communities [38, 39, 40, 41]. Meanwhile, only one marine virophage has been identified [28]. Given that mixing properties of inland water and marine environments may vary coupled with the generic lack of hyperparasitic relationships, we seek to understand how space affects virophage coexistence.

In the second chapter we consider a populations of host and virus alone. We use a spatial, individual-based model and show that spatial clustering of populations leads to increased frequency and severity of multiple infections. Varying either the viral adsorption rate or viral diffusion constant leads to the emergence of spatial clustering. This spatial clustering leads to higher abundances of viruses colocated with hosts that are already infected by viruses. These deviations in local density explain a "fattening" of the tail in the distribution of multiplicity of infection. The physical basis for these deviations are the invasion dynamics of viruses on spatial clusters of hosts.

In the third chapter we develop and analyze ODE models for the population dynamics of virophage–a virus that engages in a hyperparasitic lifestyle. The two models refer to different modes of coinfection for the virophage. In one mode, the virophage associates with the virus and the pair mutually infect the host. In the other mode, the virophage independently enters the host with coinfection initiating when a virus follows suit. We find

that both modes of coinfection lead to stable coexistence. Virophage coexistence can only occur when the host-viral dynamics are stable in the absence of the virophage. We find that both modes of coinfection reduce the viral abundances and increase host abundances as compared to the dynamics in the absence of virophage. The deviation is explained by the change in the average viral burst size resulting from fraction of lysis events involving coinfections.

In the fourth chapter we extend our analysis of the virophage system to space. Prior work has shown that the feasibility of predator-prey relationships are enhanced due to space [42]. This occurs because viral-free regions of space emerge, in which, hosts can proliferate. However, virophage rely on the presence of both hosts and viruses. Thus we show that clustering of populations lead to spatial variation in the ecologically-derived fitness of the virophage. This results in the reduction of feasibility for virophage that primarily associate with the virus.

In the fifth chapter, we present the results of new experiments investigating the infection dynamics of a giant virus on its host in both high and low levels of relative multiplicity of infection (MOI). We are able to estimate dynamical parameters such as viral burst size, host growth rate, and average latency time. In doing so we also propose a novel phenomena for understanding giant virus production: one in which latency times are shortened, rather than extended, in the case of multiple infections.

CHAPTER 2

EMERGENCE OF INCREASED FREQUENCY AND SEVERITY OF MULTIPLE INFECTIONS BY VIRUSES DUE TO SPATIAL CLUSTERING OF HOSTS

Adapted from B. P. Taylor, C. J. Penington, and J. S. Weitz, Emergence of elevated levels of multiple infections in spatial host-virus dynamics, (In review).

Broader Context: What drives rates of multiple infection in situ is largely unknown. However, natural environments tend to feature patchy distributions of populations. In this chapter we address how host-viral interactions lead to emergent clustering of populations, which, in turn, increases the frequency and severity of multiple infections. Hence, the impact of this chapter is three-fold. First, we connect viral life-history traits, such as viral adsorption rates, to the emergence of population level clustering. Next, we show that patchy distributions lead to increased frequency and severity of multiple infections. Third, this chapter provides a base line model for future work addressing how dynamical effects of multiple infections scale to the population. The Chapter also provides the baseline modeling framework for work later in the thesis where a hyperparasite virophage interacts with a giant virus and a host in a spatially explicit landscape.

Abstract

Multiple virus particles can infect a target host cell. Such multiple infections (MIs) have significant and varied ecological and evolutionary consequences for both virus and host populations. Yet, the in situ rates and drivers of MIs in virus-microbe systems remain largely unknown. Here, we develop an individual-based model (IBM) of virus-microbe dynamics to probe how spatial interactions drive the frequency and nature of MIs. In our IBMs, we identify increasingly spatially correlated clusters of viruses given sufficient decreases viral movement. We also identify increasingly spatially correlated clusters of viruses and clusters of hosts given sufficient increases in viral infectivity. The emergence of clusters is associated with an increase in multiply infected hosts as compared to expectations from an analogous mean-field model. We also observe long-tails in the distribution of the multiplicity of infection (MOI) in contrast to mean-field expectations that such events are exponentially rare. We show that increases in both the frequency and severity of MIs occur when viruses invade a cluster of uninfected microbes. We contend that population-scale enhancement of MI arises from an aggregate of invasion dynamics over a distribution of microbe cluster sizes. Our work highlights the need to consider spatially explicit interactions as a potentially key driver underlying the ecology and evolution of virus-microbe communities.

2.1 Introduction

As some of the smallest and most abundant biological entities on earth, viruses lie at the foundation of many food webs [1]. Viruses impact biogeochemical cycles by turning over an estimated 20-50% of all bacteria daily [43]. However, the interactions of individual viruses and microbial host cells are not well characterized *in situ* despite the magnitude of their aggregate effects. Only recently, the first signal of multiple infection (MI) of a host cell by viruses has been observed in natural settings [44]. This work used metagenomic analysis to identify viral genome signals from single-cell amplified genomes of microbes from oxygen-minimum marine zones. A subsequent analysis showed that MI is likely common with half of all sequenced bacterial genomes had evidence of MI [45]. A majority of coinfecting viruses in [45] derived from the same viral order, *caudovirales*. We interpret this signal as indicating high potential levels of MI by viral particles from related strains [46]. Nonetheless, it remains unknown what ecological factors drive rates of MI.

This lack of knowledge contrasts to a breadth of experimental work identifying examples of ecological and evolutionary consequences of MI. On the ecological side, life-history traits of viral infections such as the burst size [11, 12] and the latency time [10] can depend on the number of infecting viruses or multiplicity of infection (MOI). In fact, the fate of an infected cell can be MOI dependent when lysogeny is possible [13, 47]. On the evolutionary side, intracellular competition between multiply infecting viruses leads to a prisoner's dilemma scenario in game theory [20]. "Defector" viruses utilize disproportionately large amount of intracellular resources compared to "cooperating" viruses. Defectors invade populations of cooperators inadvertently resulting in a reduction of the population-wide average fitness and possibly even viral population collapse [48, 49]. In extreme cases, exploitation of resource sharing leads to the emergence of viruses that completely rely on MI to propagate such as defective interfering particles, satellite viruses, and virophage [32, 50, 31]. These experimental studies conducted in shaken flasks, i.e., a well-mixed regime, do

not provide insights as to mechanisms governing the rates of MI in complex environments.

Spatial epidemiological models have considered MI without an explicit link between cell death and viral release [51, 52]. In contrast, proposed models of viral dynamics with MI on individuals cells have focused in an immunological framework where viruses infect individual cells of a larger organism, without inclusion of explicit spatial effects [53, 54]. Prior spatial models of microbe-virus dynamics have considered plaque growth using PDEs [55, 56, 57, 58] and inidividual based models (IBMs) [59, 60] and the evolution of viral parameters using IBMs [42, 61, 62, 63]. Only [42] included MI; however, the analysis did not quantify levels of MI and instead addressed whether MIs enhance virus-microbe coexistence. The question remains: how does realistic spatial clustering of populations alter subsequent MI dynamics?

Here, we address the basis for the emergence of MI using a stochastic, spatial IBM. We quantify the frequency of MI by comparing abundances of multiply infected hosts to abundances of singly infected hosts and abundances of viruses. Additionally, we characterize the severity of MIs by tracking the distribution of MOI across hosts. We then compare levels of MI between spatial and non-spatial models across parameter ranges that vary in spatial clustering. We find that MI frequency always increases with spatial clustering, whereas single infection frequency increases or decreases depending on which populations cluster. Similarly, MI severity increases with spatial clustering as displayed by fatter tails of the MOI distribution. Finally, we show how MI is enhanced during viral invasion of host clusters. As we discuss, the inclusion of spatial dynamics gives rise to both more frequent and more severe MIs, consistent with recent genomic-based inferences of environmentally sequenced microbes [44, 45].

2.2 Methods

2.2.1 Spatial model

We develop a stochastic, spatial IBM of virus-microbe dynamics. We use an IBM over a PDE because prior work has shown demographic stochasticity increases the parameter regimes leading to pattern formation [64]. The domain is a two-dimensional, periodic square lattice where at most one host and any number of viruses can occupy a lattice point. Dynamics occur at fixed time steps given stochastic processes that include cell growth, cell death, infection, lysis, and virus decay (see Appendix). Figure 2.1 shows how multiple infections can occur during the simulation. In this example, a colocated virus infects hosts at an average adsorption rate ϕ. Multiple infections occur if another colocated virus infects the previously infected host cell before lysis. Note, in a single time step more than one virus can infect the same host cell. Infected cells lyse at a rate λ releasing a burst size β viruses into the lattice point. Infected cells act as a sink for viruses since β is assumed here to be independent of MOI.

We initiate the spatial dynamics by randomly distributing hosts and viruses given that the total initial abundances match equilibrium solutions of the analogous mean field model. Each simulation is run for 10^5 timesteps corresponding to roughly 100 days given our simulation parameters. The goal was to simulate beyond transients such that our statistics would be representative of stationary distributions. The observed steady states are robust to initiating with alternative initial conditions.

2.2.2 Mean Field ODE model

Here we present a mean-field ODE model of virus-microbe dynamics with MI:

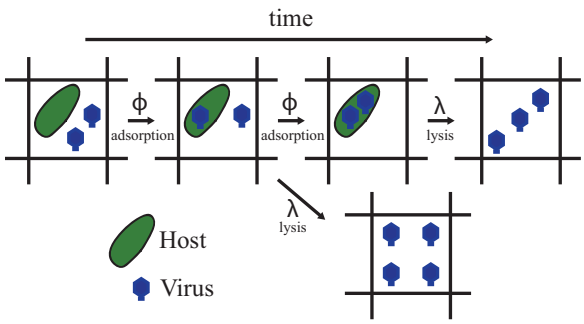


Figure 2.1: Infection dynamics in a single lattice point. Infection by a single virus occurs stochastically with adsorption rate, ϕ. The total rate of infection depends on the local viral abundance, ϕV_{local}. Infected cells lyse with rate λ independent of MOI. Infected cells can be multiply infected if another infection event occurs before lysis. Lysis removes the host cell and replaces it with new viruses according to a fixed burst size independent of MOI. The burst size in this cartoon is 3 viruses for graphical convenience. We use a burst size of β =20 in our models.

$$
\begin{aligned}
\dot{H} &= \overbrace{rH\left(1-\frac{H+\sum_j I_j}{K}\right)}^{\text{logistic growth}} - dH - \overbrace{\phi V H}^{\text{infection}} \\
\dot{I}_1 &= \overbrace{\phi V(H-I_1)}^{\text{infection}} - \overbrace{(d+\lambda)I_1}^{\text{death \& lysis}} \\
\dot{I}_2 &= \overbrace{\phi V(I_1-I_2)}^{\text{infection}} - \overbrace{(d+\lambda)I_2}^{\text{death \& lysis}} \\
&\vdots \\
\dot{V} &= \overbrace{\beta\lambda\sum_j I_j}^{\text{lysis}} - \overbrace{\phi\left(H+\sum_j I_j\right)}^{\text{infection}} - \overbrace{mV}^{\text{decay}}
\end{aligned}
\tag{2.1}
$$

where H tracks uninfected hosts, I_j tracks infected hosts with MOI=j, and V tracks the viruses. We are interested in the dynamics of multiply infected cells, $C = \sum_{j\geq 2} I_j$ which follow:

$$
\dot{C} = \overbrace{\phi V I_1}^{\text{infection}} - \overbrace{(d+\lambda)C}^{\text{death \& lysis}}. \tag{2.2}
$$

The parameters are the same between the spatial and mean field model. Differences between the two models result from spatial degrees of freedom and stochasticity. The steady-state solutions to the mean-field model can be solved exactly (see Appendix). The model assumes that both burst size and lysis rate are independent of MI.

We can obtain relevant MI statistics by solving the mean field model at equilibrium. From (2.2), we expect:

$$C^* = \frac{\phi I_1^* V^*}{(d+\lambda)} \tag{2.3}$$

where * denotes abundances at steady-state. Similarly, solving $\dot{I}_1 = 0$ in (2.1) yields:

$$I_1^* = \frac{\phi H^* V^*}{(d+\lambda)}. \tag{2.4}$$

Solving (2.1) at equilibrium for any of the multiply infected host classes leads to a geometric sequence for the MOI distribution:

$$\frac{I_j^*}{I_{j-1}^*} = \frac{\phi V^*}{\phi V^* + d + \lambda}. \tag{2.5}$$

These derived results allow identifying deviations of the spatial model from the mean field expectation based on observed abundances.

2.3 Results

2.3.1 Spatial clustering emerges as viral adsorption and diffusion vary

We independently vary the adsorption rate ϕ and the viral diffusion constant D_v leading to emergent clustering in the spatial dynamics. The top two rows of Figure 2.2 show snapshots of the host and viral spatial dynamics respectively while only varying the adsorption rate, ϕ. Spatial clustering increases with increasing ϕ (columns moving left to right) as quantified by radial pair correlation profiles shown in the bottom row of Figure 2.2. These profiles

display the ratio of the average number of hosts (viruses) located some distance away from each host (virus) over the expectation from random dispersal over the domain [65]. Values greater than 1 indicate increased frequency of hosts at a given distance. We define clustering to occur when the profiles are greater than a threshold value of 1.1 (dashed line). We chose this threshold as it is approximately the maximally observed 99% confidence interval when comparing to simulated ensembles of random dispersal. Low ϕ dynamics lead to cluster profiles indistinguishable from random dispersal of hosts and viruses (left column). As ϕ increases small clusters of hosts increase in frequency whereas the viral cluster profile is indistinguishable from random dispersal (middle column). High values of ϕ, where the dynamics are still robust against stochastic extinction events throughout our simulations, feature spatial clustering of both hosts and viruses (right column). Thus as ϕ varies we observe spatial dynamics feature no clustering, host clustering alone, then both host and viral clustering. We also varied the viral diffusion constant, D_v, which lead to increased viral clustering alone as D_v decreased (see Appendix). Interestingly, varying D_v did not lead to host clustering over the explored range. Viral clustering alone leads to temporary virus-free domains where hosts can locally reproduce to the local carrying capacity.

Independently increasing ϕ or decreasing D_v leads to increased clustering. The top row of Figure 2.3 shows this increase in clustering in terms of cluster widths. We define cluster widths as the maximal distance where the pair correlation profile exceeds the threshold. The range of parameters are colored according to whether the host, the virus, or both populations have nonzero cluster widths. Note, all data points from the simulations are averages over the last 10^4 timesteps and over 10 replicates. The corresponding cluster widths for our prior examples are marked by black lines in the top two rows of Figure 2.2.

2.3.2 Effects of spatial clustering: increased relative rates of multiple infection

Here, we explore the effects of clustering on population densities and MI. To begin, we find that increased clustering leads to modest deviations in virus and host abundances from

mean field expectations. The 2nd row of Figure 2.3 shows how most population abundances decrease from the mean-field expectation as a result of increased clustering. Only the uninfected host abundance increases compared to mean-field expectation as a result clustering when varying D_v. In that case, a smaller value in D_v limits viral movement leading to clustering of the viruses alone. Accordingly, the space between viral clusters acts as temporary domains where the hosts can grow uninhibited.

Spatial clustering leads to modest deviations in rates of infection that depend on the form of clustering observed. We estimated the time-averaged infection rate in the simulations by counting the number of new infections over 240 time-steps corresponding to 6 hours. We define the relative infection as the ratio of our observations to the mean-field expected infection rate of $\phi H^s V^s$ where the superscript "s" denotes abundance observed from the simulations. The 3rd row of Figure 2.3 plots the relative infection rate. Points above (below) the black line indicate larger (smaller) rates of initial infection as compared to the expectation based on the observed uninfected host abundances and the viral abundances. Increased clustering due to varying ϕ leads to increased relative infection rates. Meanwhile increased clustering due to varying D_v leads to decreased relative infection rates. This latter deviation is associated with negative spatial correlation between hosts and viruses due to increased fractions of the host population existing in temporary virus-free domains.

In contrast, clustering leads to significant increases in the MI frequency regardless of the form of clustering. The bottom row of Figure 2.3 shows the relative MI frequency previously defined in the methods section and reiterated in the caption of Figure 2.3. Points above the line indicate larger abundances of multiply infected hosts as compared to the observed singly infected host abundances and the viral abundances.

2.3.3 Effects of spatial clustering: MOI distributions with fatter tails

In this section, we show spatial clustering increases the severity of MIs as described by the MOI distribution – the abundance of hosts of increasing MOI. The MOI distribution is

relevant when life-history traits can be MOI dependent (e.g., increased burst sizes, longer latency times). Our analysis gives a baseline understanding of the relevance of MOI skewing, though we do not explicitly modify life history traits as a function of MOI. The top rows of Figure 2.5 show the observed MOI distributions as compared to the previously derived mean-field expectation (black line) as we vary ϕ. We normalize the distributions by setting the density of singly infected hosts to 1. The observed MOI distributions match the mean-field geometric sequence for low ϕ. Spatial clustering leads to an increase in the mass of the tail of the MOI distribution, i.e., more hosts are infected by more viruses as compared to mean field. This deviation occurs continuously across parameter space as evidenced by the slight deviation present for weakly clustered dynamics (middle column Figure 2.5). The analogous plots when varying D_v are shown in the Appendix.

The MOI distributions have relatively "fat" tails because more viruses are colocated with hosts of increasing MOI. We address the role of density fluctuations of viruses by calculating viral probability distributions (VPDs). Figure 2.4 shows the construction of VPDs by counting up the local viral population conditioned on being colocated with a host of a certain MOI. The bottom row of Figure 2.5 shows the probability distributions of observing a number of external viruses in lattice points that contain hosts with a specific MOI. For clarity we only show these VPDs up to MOI=6. Randomly distributing hosts and viruses across the domain leads to Poisson distributed VPDs (black lines) parametrized by the observed viral density. The observed VPDs match the Poisson distribution for all MOI in the low ϕ. Whereas for all other cases, the observed VPDs deviate from the Poisson distribution by skewing to the right, i.e., there are more viruses colocated with high MOI hosts than expected in the mean field theory. This skewing is more pronounced as clustering increases.The analogous plots when varying D_v are shown in the Appendix.

The viral distributions dictate the rate of flow between MOI types in the dynamics. The rate of viral infection in a lattice site linearly depends on the number of colocated external viruses. Here, we propose to adapt the mean-field approach, taking into account how local

viral densities alter the dynamics. For example, the infection dynamics should follow

$$\dot{I}_j \;=\; \phi(V_{j-1}I_{j-1} - V_jI_j) - (d+\lambda)I_j. \tag{2.6}$$

At equilibrium, this yields a sequence of multiplicative factors for the abundances of increasing MOI hosts:

$$\frac{I_j^s}{I_{j-1}^s} \;=\; \frac{\phi\langle V_{j-1}\rangle}{\phi\langle V_j\rangle + d + \lambda}. \tag{2.7}$$

where the superscript "s" refers to the observed abundances from the spatial model that we conjecture to follow this relationship. We have replaced V_j^* and V_{j-1}^* with $\langle V_j\rangle$ and $\langle V_{j-1}\rangle$ respectively, i.e., the means of the corresponding VPDs from the spatial IBM. The MOI distribution built from these scaling factors (red lines) matches the observed distributions in Figure 2.5.

2.3.4 Cluster invasion dynamics skew VPDs

We now show that viruses increasingly colocate with high MOI infected hosts during the viral invasion dynamics of a cluster of initially uninfected host cells. These invasion events become relatively more frequent with increased clustering and, in turn, further skew the statistics of the full system. For example, the top two rows of Figure 2.6 show a series of snapshot of the spatial dynamics with an adsorption rate $10^{0.2} \approx 1.6$ times the maximum of the range explored in Figure 2.3. The typical dynamics at this parameter value oscillate wildly ultimately leading to stochastic extinction of at least one population within our simulation period. Here, most of the hosts and viruses are grouped in dense compact clusters (left frame). A single virus diffuses into a cluster initiating infection (middle frame). Note the source of this virus is not from the larger, discernible viral cluster in the left column. Rather, this virus remains from a previous cluster invasion event that occurred nearer. After lysis of the single infection event a cascade of infections leads to a wave of infections

across the cluster (right frame). A large viral cluster remains after clearing the host cluster. Survival of the virus population relies on diffusing to and infecting a nearby growing host cluster before decay.

In this example, the statistics and dynamics of the full system are completely determined by the growth and subsequent invasion of these clusters. To explore the effect of these dynamics on our MI statistics we simulate invasions on single host clusters. We consider maximally dense clusters of hosts of increasing radii and initiate a lytic event at the center. We utilize invasion simulations with immobile hosts that do not reproduce or diffuse to examine solely the effect of lysis and viral diffusion on MI. Figure 2.6 shows the time-averaged dynamics of the VPD for invasions of clusters with increasing radii. Each figure shows the VPDs of a specific MOI. For all MOI, the VPD skew to the right as the radius of the cluster increases, consistent with the results from the full stochastic IBM.

2.4 Discussion

In this paper, we demonstrated that both the frequency and severity of multiple infections (MIs) of microbial hosts by viruses increases due to spatial clustering. We identified the increase of MI frequency by comparing observed abundances of multiply infected host cells to abundances of singly infect host cells and viruses. The increased severity of MI was characterized by fatter tails in the multiplicity of infection (MOI) distribution. This fatter tail arises from positive skewing in the distribution of external viruses colocated with hosts of higher MOI types. Finally, we argued that invasion of larger host clusters leads to the skewed VPDs.

Part of the motivation for studying MI rates is that MIs can alter the dynamics of individual infections and, in turn, the entire population. While our models do not include MOI dependent parameters, our results do provide a baseline to compare future models where these feedbacks are included. For example, intracellular competition amongst multiply infecting viruses is more likely given spatially clustered dynamics. Such an increase in in-

tracellular competition may lead to evolutionary conflict between viruses, e.g., a prisoner's dilemma [20] and even to the extinction of a population [66]. In addition, the emergence of MI may also indicate when mechanisms to prevent secondary infections, like superinfection exclusion, should evolve [67]. Similarly, our model may also have implications for understanding the long-standing puzzle of persistence of multipartite viruses that require a high MOI for successful propagation [68].

While the patterns formed in our system are due to localized growth and limited dispersal, high density regions can occur by other means. The observed deviations between models arise from spatial correlations between host and viral types and are thus robust to specifics of how clustering forms. Hence, MIs could play a major role in shaping the population dynamics across a wide-range of patchy communities (e.g., in biofilms [69], during ocean blooms [70], resulting from chemotaxis [71] and resulting from turbulence [72], and standing patchiness [73, 74]). Furthermore, our observation of viral clustering in the absence of host clustering suggests that MIs may play a major role even in environments without observed microbial patchiness. Our simulations suggest this is more likely to occur for environments with a high host-viral ratio. Overall, observing rates of MI is of major empirical importance for understanding virus-microbe dynamics *in situ*. While recent work has demonstrated the existence of MIs in a targeted marine microbe [44] and within sequenced genomes [45] quantitative measurements of the frequencies are lacking.

In summary, this paper sheds light on how multiple infections emerge from population-scale dynamics taking place in spatially explicit domains. MOI dependent life-history traits can then act to modify subsequent population dynamics. It remains a question as to whether or not these kinds of feedback amplify or reduce MIs. In particular, MI allows for direct competition of viruses via shared resources inside the host. This complicates viral evolution as exploiting the host must be counterbalanced by exploitation from further viruses. Thus the effect of MI on viral evolution at a population level is relatively unexplored. The ubiquity of spatial clustering in natural environments suggests that increased attention on

MI is necessary to understand the eco-evolutionary dynamics of the microbes and their viruses.

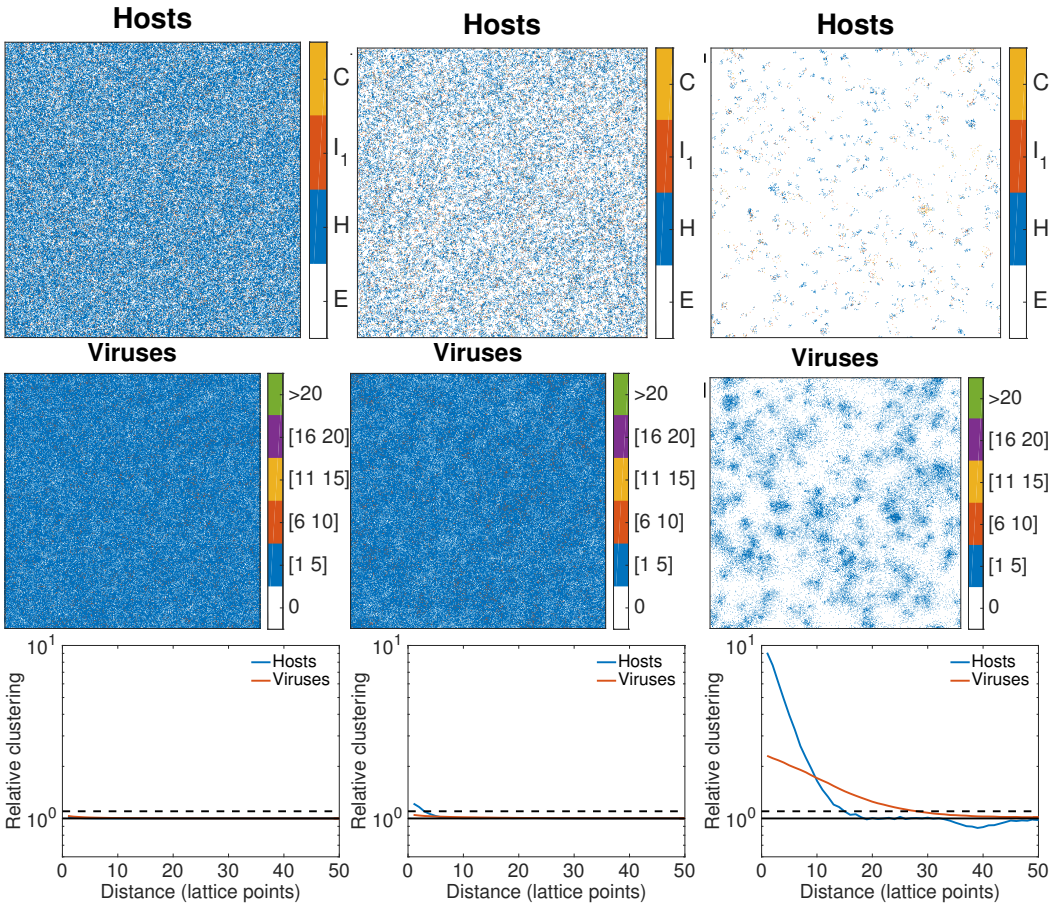


Figure 2.2: Spatial clustering increases with increasing adsorption. The color legend for hosts refers to multiply infected hosts (C), singly infected hosts (I_1), uninfected hosts (H) and empty lattice sites (E). The color legend for viruses refers to the number of viruses located at each lattice point. (Left Column) Low adsorption, $\phi = 10^{-8.4}$ (ml/hr). (Middle Column) increased adsorption, $\phi = 10^{-8.0}$ (ml/hr) (Right Column) High stable value of adsorption, $\phi = 10^{-7.1}$ (ml/hr). Rows correspond to (Top) distribution of hosts, (Middle) distribution of viruses and (Bottom) radial pair correlation profile of hosts. The dotted line at $y = 1.1$ approximates the 99% confidence interval of the pair correlation profile when hosts and viruses are randomly dispersed. We use the intersection of this threshold line and the observed pair correlation profiles to define the cluster widths. When clustering occurs the corresponding cluster widths are plotted as black lines outside the top left corner of each of the spatial distribution plots. See appendix for other parameter values.

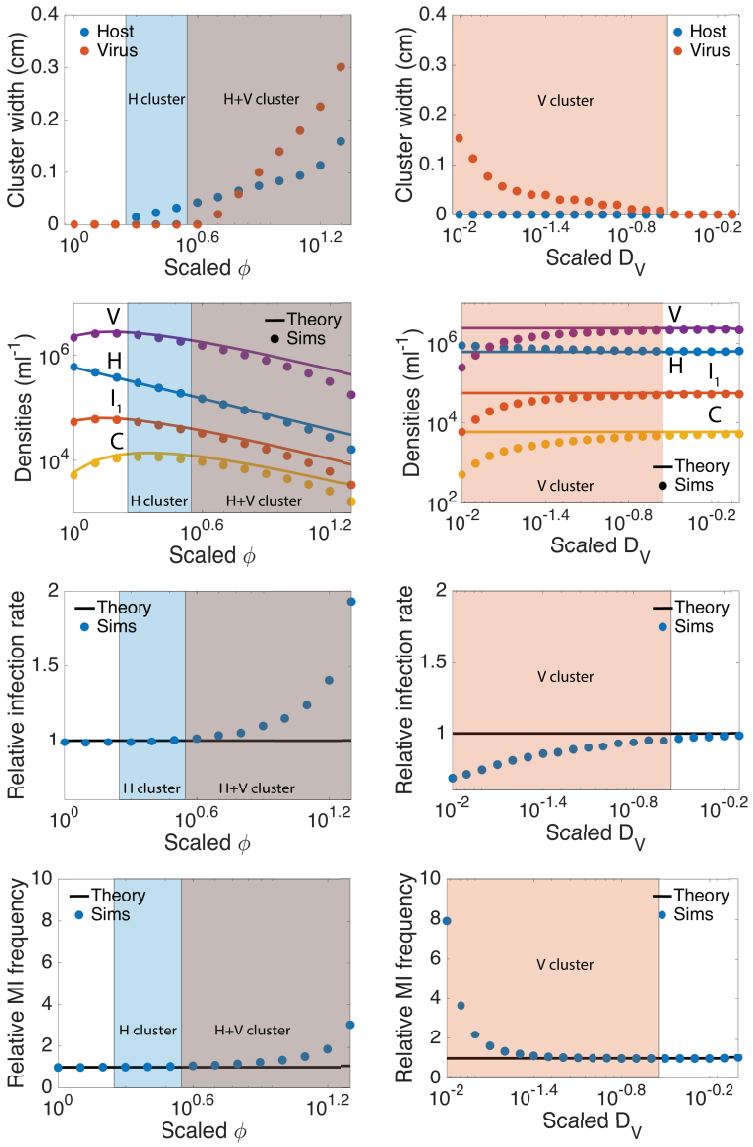


Figure 2.3: Emergence of clustering and its effects on densities and MI. (Left column) Results obtained by varying the adsorption rate. (Right column) results obtained by varying the viral diffusion constant. The x-axes refer to scaling from reference parameter values of $\phi = 10^{-8.4}$ (ml/hr) and $D_v = 2.04 \times 10^{-4}$ (cm^2/hr). All points are time and ensemble averages over the last 10^4 time steps and 10 replicate simulations. (First row) Cluster widths determined from non-zero x-value of the intersection between the pair correlation profiles and chosen threshold line. The transparent patches throughout indicate parameter values featuring clustering by hosts (blue), viruses (orange), or both (blue and orange overlay). (Second row) Population abundances where the line corresponds to solutions of the analogous non-spatial ODE model. (Third row) Relative rate of initial infections defined as the ratio of time-averaged new infections over 240 time-steps over the mean-field expected infection rate, $\phi H^s V^s$ where "s" denotes abundances observed from the simulations. The dotted line is unity. Points above (below) the dotted line indicate more (fewer) initial infections given observed abundances of uninfected hosts and viruses. (Fourth row) Relative frequency of MIs rate quantified by $C^s(d+\lambda)/(\phi I_1^s V^s)$ where the superscript "s" denotes abundance observed from the simulations. The dotted line is the mean-field expectation of unity. Points above (below) the dotted line indicate more (fewer) multiply infected hosts given observed abundances of singly infected hosts and viruses.

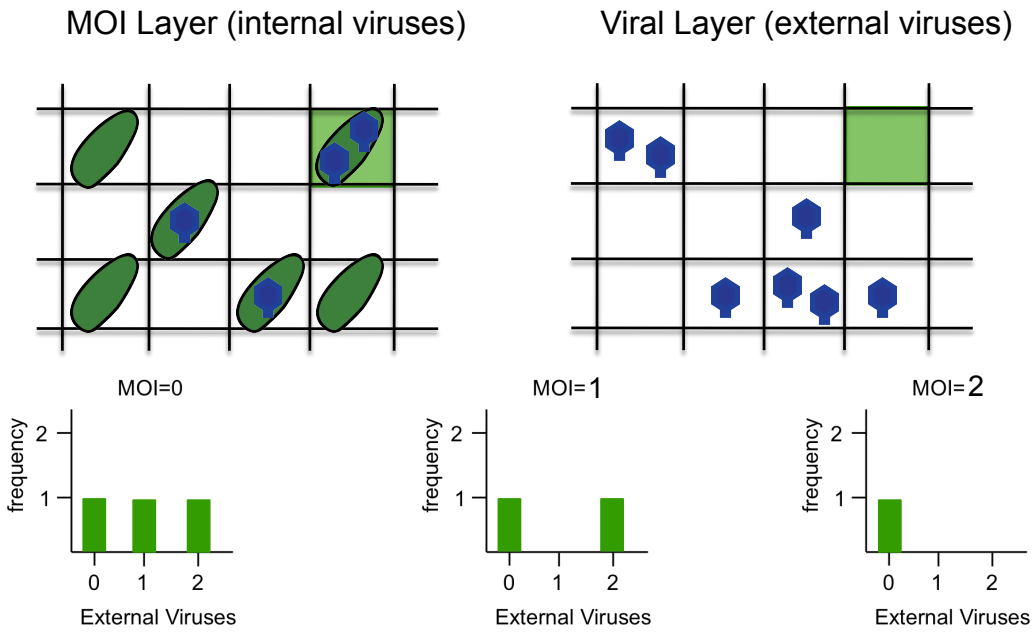


Figure 2.4: Example state and associated viral probability distributions (VPDs). VPDs are constructed from histograms of external virus populations conditioned on the MOI state of the host. The highlighted lattice point represents conditioning on MOI=2 and constructing the histogram over the respective lattice points.

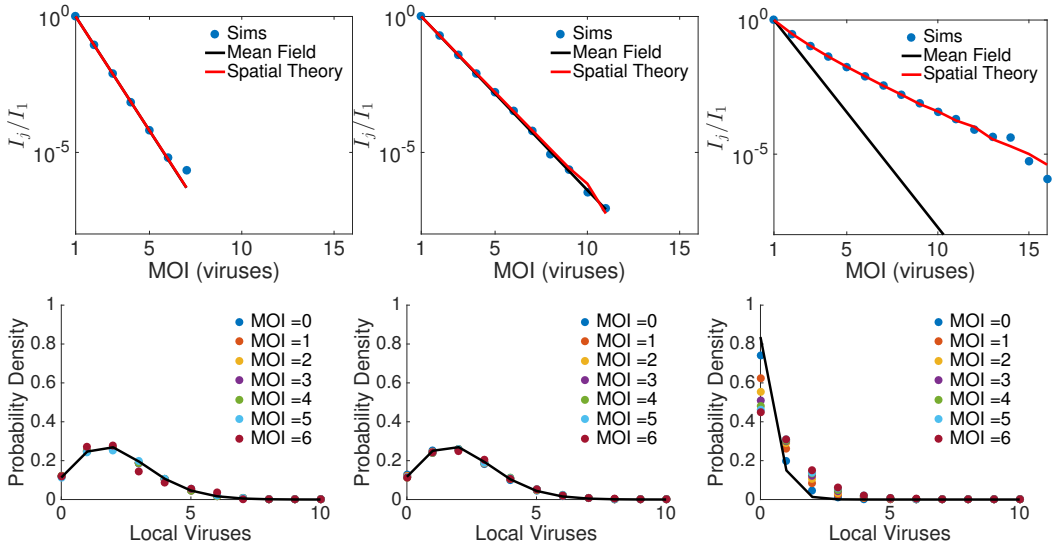


Figure 2.5: MOI distributions arising from local densities of viruses and underlying viral probability distributions (VPDs). VPDs quantify the local densities of viruses as normalized histograms of viral abundances colocated with host of a specified MOI. Predictions from spatial theory obtained using Eqn. 2.7. (Top) MOI distributions and (Bottom) normalized distributions of colocated viruses conditioned on MOI host for (Left Column) Low adsorption, $\phi = 10^{-8.4}$ (ml/hr), (Middle Column) increased adsorption, $\phi = 10^{-8.0}$ (ml/hr), and (Right Column) high stable value of adsorption, $\phi = 10^{-7.1}$ (ml/hr).Other parameters are the same as in Figure 2.2

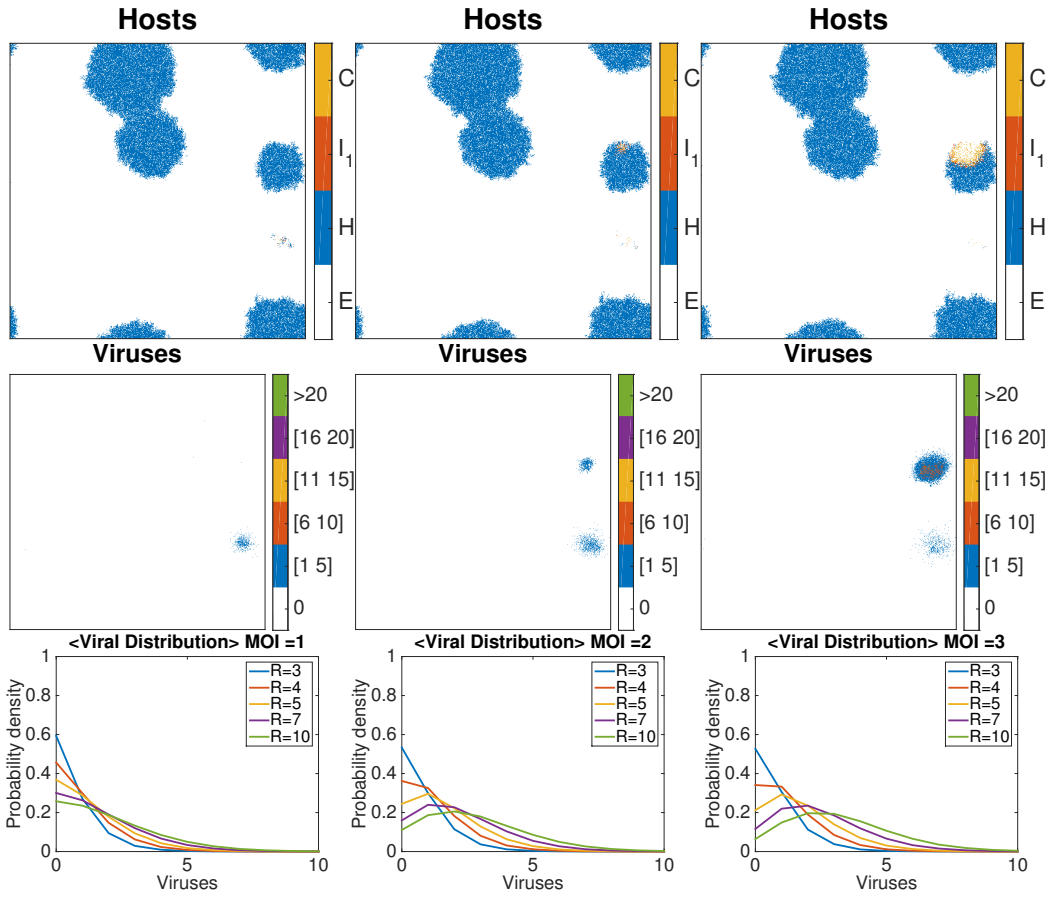


Figure 2.6: Virus infections spread wave-like across host clusters, enhancing MI. Snapshots of (Top) host and (Middle) viral dynamics when adsorption is increased such that unstable oscillations arise due to clustering ($\phi = 10^{-6.9}$ ml/hr). The dynamics feature the invasion of a host cluster by viruses over 50 times steps. The color legend for hosts refers to: C-multiply Infected hosts, I_1- singly infected hosts, H-uninfected hosts, E- empty lattice points. The color legend for viruses refers to the number of viruses located at each lattice point. (Bottom) Time averaged viral distributions for invasions of disc-shaped clusters of increasing radius (R) in units of lattice points. Each figure is conditioned on a different host MOI. The viral adsorption rate is $\phi = 10^{-7.1}$ (ml/hr) for the invasion dynamics.

CHAPTER 3

THE VIRUS OF MY VIRUS IS MY FRIEND: ECOLOGICAL EFFECTS OF VIROPHAGE WITH ALTERNATIVE MODES OF COINFECTION

Adapted from B. P. Taylor, M. H. Cortez, and J. S. Weitz, The virus of my virus is my friend: ecological effects of virophage with alternative modes of coinfection, J. Theor. Biol. 354, 124-136 (2014).

Broader Context: The effect of hyper-parasitism is an understudied area of ecology. The discovery of virophage 10 years ago prompts addressing the effect of hyper-parasitism in a viral setting. Unique to virophage is the potential for two different modes of coinfection: associating with the virus first or associating with the host first. This chapter provides the first biologically-motivated model of virophage dynamics. Our analysis explains why hyperparasitism is not observed as frequently as predator-prey relationships: it is necessary for the host and virus to be able to coexist alone in order for the virophage to also coexist. We also show, on the basis of feasible parameter space, that the virophage that associate with the host first are likely to be more prevalent than the other virophage type. This chapter also sets the underlying nonlinear model to be utilized later in this thesis when investigating spatially explicit dynamics of virophage, viruses, and hosts.

Abstract

Virophages are viruses that rely on the replication machinery of other viruses to reproduce within eukaryotic hosts. Two different modes of coinfection have been posited based on empirical observations. In one mode, the virophage and virus enter the host independently. In the other mode, the virophage adheres to the virus so both virophage and virus enter the host together. Here we ask: what are the ecological effects of these different modes of coinfection? In particular, what ecological effects are common to both infection modes, and what are the differences particular to each mode? We develop a pair of biophysically motivated ODE models of viral-host population dynamics, corresponding to dynamics arising from each mode of infection. We find both modes of coinfection allow for the coexistence of the virophage, virus, and host either at a stable fixed point or through cyclical dynamics. In both models, virophage tend to be the most abundant population and their presence always reduces the viral abundance and increases the host abundance. However, we do find qualitative differences between models. For example, via extensive sampling of biologically relevant parameter space, we only observe bistability when the virophage and virus enter the host together. We discuss how such differences may be leveraged to help identify modes of infection in natural environments from population level data.

3.1 Introduction

Virophages are recently discovered viruses of viruses [31, 28]. To reproduce, a virophage must infect a eukaryotic host that is also infected by a larger virus [31]. These larger viruses, hereafter referred to as viruses, are classified as NucleoCytoplasmic Large DNA Viruses and have comparable physical sizes and genome lengths to small bacteria [75]. These larger viruses require the host to reproduce; however, their relatively large genomes encode for their own transcriptional machinery and part of their translational machinery, allowing the virus to reproduce within the host in a cytoplasmic structure of viral origin termed the "viral factory" [27]. When also present, the virophage is thought to utilize the transcriptional machinery of viral origin and reproduce within the viral factory. The virophage genome is much smaller than the genome of the virus and does not encode for any of the constituent parts of the viral factory. Hence, virophages reproduce obligately through coinfection. The virophage serves a parasitic role to the virus as viral burst sizes are greatly reduced during coinfection [31, 76].

Virophages are continually being discovered and appear to be widespread biological entities in clinical and environmental settings. The first discovered virophage, termed Sputnik, was isolated from a virus, mamavirus, that was extracted from the water in cooling towers in Paris, France [31]. A later discovered strain of Sputnik, termed Sputnik2, is associated with mamavirus-like Lentillevirus and shares the host *Acanthamoeba polyphaga*, which is a causative agent of the human eye disease keratitis [77]. More recently, a third strain of Sputnik was discovered along with evidence that all strains could associate with many more viral strains than previously thought [78]. A different but related virophage, termed Organic Lake Virophage (OLV), was discovered from environmental sequencing data obtained from a hyper-salinic Antarctic lake and is associated with an algal host that undergoes yearly bloom cycles [38]. The first discovered marine virophage, Mavirus, is associated with the bacterivorous host, *Cafeteria roenbergensis*, which is endemic among

the global oceans [28]. These last two examples suggest that newly discovered virophages may have global implications on algal blooms and nutrient cycles [79]. In fact, a genomic study suggests that undiscovered virophages exist in dozens of more locations including at different depths in oceans and lakes across the globe and within humans and other animals [39].

Among the discovered virophages, two different primary means for coinfection seem plausible. In one mode, which we call the independent entry mode, the virophage and virus independently enter the host cell. In the other mode, which we term the paired entry mode, the virophage entangles with the virus and coinfection occurs when the composite enters the host. The paired entry mode of coinfection is thought to be the utilized by Sputnik strains [76]. In addition, the paired-entry mode of coinfection, to our knowledge, has strong indirect support, e.g., images show virophage grouped around viruses during viral production suggesting an affinity [76]. Images of virophage and virus present in the same phagocytic vacuole after coinfection serve as further evidence [76]. Additionally, there is a hypothesized structural basis for virophage-virus entanglement. The mamavirus, a virus associated with Sputnik strains, is coated with long, tendrils that likely function to induce phagocytosis. Experimental tests suggest these fibers are coated with peptidoglycan. This coating is hypothesized to promote viral mimicry of the bacterial prey of the host amoeba [80]. Mushroom-like fibers coat the capsid exterior of the virophage Sputnik [81]. The function of these fibers are unknown, but it is hypothesized they interact with mamavirus fibers to promote associating into a composite [76]. In accordance with this hypothesis, Sputnik is unable to reproduce in mixed cultures with bald forms of mamavirus-like strains [82].

The two modes for coinfection are pieces of a larger virophage infection process. The entire coinfection processes for the independent entry and the paired entry modes are shown in Figures 3.1a and 3.1b, respectively. The post-coinfection dynamics are considered equivalent between modes (steps 4-6). Here, the viral core and virophage genome (of the virus

and virophage, respectively), separate from their capsids (step 4). Hereafter, we refer to viruses and virophage collectively as viral particles. Note that this step has been experimentally observed for the virus but not for the virophage [27]. The viral factory originates from the viral core, which contains the viral genome. (step 5)][83, 27]. Capsids form on the exterior of the viral factory and fully formed viral particles remain in the host cytoplasm until host lysis occurs. Lysis typically occurs at about 16 hours post-infection for Sputnik [76].

In this paper, we use theoretical models to explore how the ecological dynamics of the host, virus, and virophage populations depend on the biophysical mechanism of coinfection. Our models correspond to the independent entry and paired entry mechanisms above. A particularly important mechanistic difference between our models and two other models of virophage dynamics in the literature, is that we explicitly model the population dynamics of the viral particles in the environment. One model treated the virophage as a predator of the virus and modeled virophage growth as host independent (illustrated in Figure 3.1)[38]. However, virophages require both host and virus for reproduction. The other model borrows from epidemiological theory by modeling the spread of viruses and virophage through direct-contact between hosts, i.e., it does not model free virus or virophage in the environment (illustrated in Figure 3.1)[84]. We note that infection dynamics from models of direct and indirect disease transmission can coincide when viral dynamics (e.g., degradation) in the environment are very fast [85]; however, we are unaware of experimental evidence to suggest this is the case.

In the rest of the paper, we first present our mathematical models for each mode of coinfection. Next, we demonstrate that stable and cyclical coexistence occurs between the virus, virophage, and host in each mode. In both models, we find that virophage coexistence results in a reduction of viral abundance and an increase in host abundance. We then derive an effective theory of host-viral interactions that accounts for this virophage-mediated shift in population levels. Finally, we identify differences in coexistence between two modes

that may be leveraged in future efforts to identify the infection mode from population level data.

3.2 Methods

3.2.1 General modeling framework

For both modes of coinfection, we explicitly model the density of viruses and virophage in the environment (step 1 in Figures 3.1a,b). The units for these densities are ml^{-1}. Intracellular dynamics (lysis) are assumed to be instantaneous once infection occurs. This is akin to step 3 pointing directly to step 1. We assume a well-mixed system and assume the rates of contact follow mass action kinetics. The host and viral dynamics are modeled from an adapted Lotka-Volterra framework. Further assumptions specific to each model are discussed below. Note that consideration of delays between infection and lysis have been previously analyzed in the viral modeling literature and would be an important area of future exploration in this context [86, 87].

IEM: independent entry model of virophage and virus

In the independent entry model (IEM), the virus and virophage independently enter the host. Coinfection occurs when a virus enters a host in which the virophage previously entered. We model the dynamics between the host (H), virus (V), virophage (P), and the host with an internal form of the virophage (H_p), hereafter referred to as an infected host,

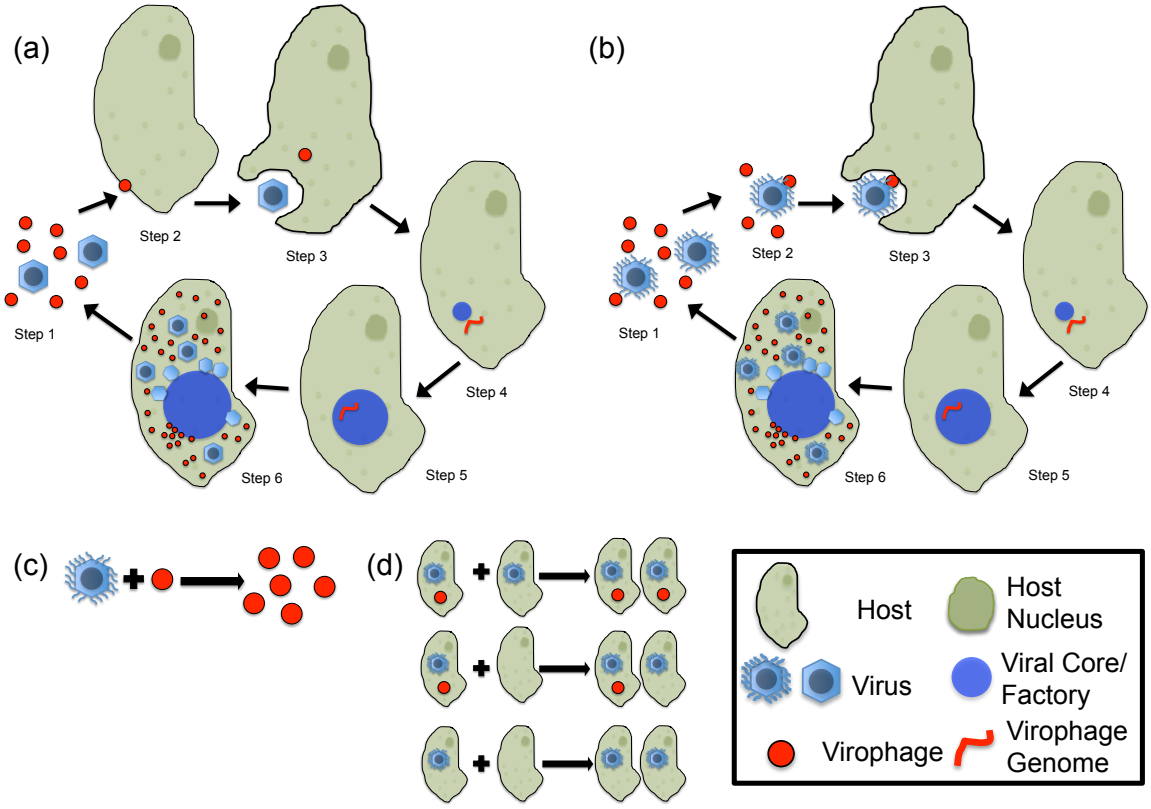


Figure 3.1: Stages during the virophage coinfection cycle in alternative mathematical models. (a) Independent entry mode, developed here: Step 1: free virus and virophage in the environment following host lysis. Step 2: free virophage in the environment enters the host, note the host nucleus is shown as an internal large, darker green circle. Step 3: free virus enters a host that previously engulfed a virophage. Step 4: the viral particles lose their capsids. Step 5: the virophage genome enters the viral factory (large blue circle) which expands as viral particle genome replication occurs internally. Step 6: fully formed viral particles bud from the viral factory and remain in the host cytoplasm until host lysis occurs. (b) Paired entry mode, developed here, only steps 2 and 3 differ from the previous model. Step 2: virophage attaches to virus to form a composite in the environment. Step 3: the composite enters the host causing coinfection. (c) Mechanism of virophage reproduction from a previous model: virophage reproduce via the infection and lysis of the virus in the absence of a host [38]. (d) Reproduction of virophage and virus in a direct contact model where free viral particles in the environment are not modeled [84].

see Figure 3.1a, steps 1-3. The full model is:

$$
\begin{aligned}
\frac{dH}{dt} &= \overbrace{H\left[b - d\left(1 + \frac{H + H_p}{K}\right)\right] + (1-\rho)bH_p}^{\text{host growth}} - \overbrace{(\phi_p P + \phi_v V)H,}^{\text{infection (virophage and virus)}} \\
\frac{dH_p}{dt} &= \overbrace{H_p\left[\rho b - d\left(1 + \frac{H + H_p}{K}\right)\right]}^{\text{infected host growth}} + \overbrace{\phi_p P H - \phi_v V H_p,}^{\text{infection (virophage and virus)}} \\
\frac{dV}{dt} &= \overbrace{(\beta_v H + \beta_{vp} H_p)\phi_v V}^{\text{virus production (lysis)}} - \overbrace{m_v V,}^{\text{virus decay}} \\
\frac{dP}{dt} &= \overbrace{\beta_p \phi_v V H_p}^{\text{virophage production (lysis)}} - \overbrace{\phi_p P H}^{\text{virophage infection}} - \overbrace{m_p P,}^{\text{virophage decay}}
\end{aligned}
\tag{3.1}
$$

where b and d are the density-independent host birth and death rates, respectively, and ρ is the fraction of infected host offspring that remain infected after reproduction. K is the host density at which the total death rate is twice that of the intrinsic death rate, ϕ_v is the absorption rate between virus and host, and ϕ_p is the host absorption rate of virophage. The virus and virophage decay with rates m_v and m_p, respectively. The burst size of the virus is β_{vp} during virophage coinfection and β_v otherwise, while β_p is the burst size for virophage. We assume virophage and virus burst sizes during coinfection are linearly dependent on β_v such that $\beta_p = \rho_p \beta_v$ and $\beta_{vp} = \rho_{vp}\beta_v$. We assume virophage can not enter a host with a virophage already present. We assume the virophage does not decay within the host. We also assume the host pays no cost and gains no direct benefit while carrying the virophage.

PEM: paired entry model of virophage and virus

In the paired entry model (PEM), coinfection occurs when a virophage attaches to a virus in the environment and the virophage-virus composite later enters the host. We model the population dynamics of the host (H), virus (V), virophage (P), and the virophage-virus composite (V_p), as shown in steps 1-3 of figure 3.1b. The full model is:

$$
\begin{aligned}
\frac{dH}{dt} &= \overbrace{H\left[b - d\left(1 + \frac{H}{K}\right)\right]}^{\text{host growth}} - \overbrace{\phi_v(V + V_p)H,}^{\text{infection (virus and composite)}} \\
\frac{dV}{dt} &= \overbrace{(\beta_v V + \beta_{vp} V_p)\phi_v H}^{\text{virus production (lysis)}} - \overbrace{\phi_{vp} V P}^{\text{virophage adhesion}} + \overbrace{m_p V_p}^{\text{virophage decay}} - \overbrace{m_v V,}^{\text{virus decay}} \\
\frac{dV_p}{dt} &= \overbrace{\phi_{vp} V P}^{\text{composite formation}} + \overbrace{\beta_i \phi_v V_p H}^{\text{composite burst (lysis)}} - \overbrace{(m_p + m_v)V_p,}^{\text{decay (virus and virophage)}} \\
\frac{dP}{dt} &= \overbrace{\beta_p \phi_v V_p H}^{\text{virophage production (lysis)}} - \overbrace{\phi_{vp} V P}^{\text{composite formation}} + \overbrace{m_v V_p}^{\text{virus decay}} - \overbrace{m_p P,}^{\text{virophage decay}}
\end{aligned}
\tag{3.2}
$$

where ϕ_{vp} is the rate of entanglement between virus and virophage and β_i is the virophage-virus composite burst size. We assume the composite burst size is linearly dependent on β_v such that $\beta_i = \rho_i \beta_v$. The rest of the parameters have the same meaning as in the IEM. We assume only one virophage can entangle with a virus and, once entangled, either the virus or virophage can independently decay leaving the virophage or virus free, respectively. We assume that some viruses and virophage emerge as composites after lysis. We do not include the incorporation of the virophage (provirophage) into the viral DNA as observed between one strain of the virophage Sputnik-2 and one strain of the virus Lentillevirus [88].

3.2.2 Biophysical parameters

We obtained reference values for the model parameters either from the literature, from derivations based on first principles, or through personal communication (with Matthias Fischer). We used the Mavirus virophage system as a reference for our IEM parameters. Mavirus has been observed independently entering the host; however, the exact mechanism is not well understood and can not be definitively identified as IEM as we modeled here [28]. The PEM parameters are in reference to the Sputnik-Mamavirus-*Acanthamoeba* system.

The reference values for the IEM and the PEM are shown in Tables 3.1 and 3.2, re-

Table 3.1: IEM reference parameters. The $\star$ denotes parameters shared between models. The values of the shared parameters may differ between models as they refer to different sets of organisms. When a value is stated for a parameter in the literature for only one system that value is used as the reference for both models as long as it is reasonable. The parameters with "=" show the relation between the parameters present in the model and the free parameters used for sampling. Personal communication with M. Fischer (MPI-Heidelberg).

Symbol	Meaning	Value	units	Reference
$\star K$	carrying capacity	4.0×10^6	$\frac{host}{ml}$	[28]
$\star b$	host birth rate	2.7	day^{-1}	personal communication
$\star d$	host death rate	1.4	day^{-1}	assumed; see B
$\star m_v$	viral decay rate	$6.3 * 10^{-2}$	day^{-1}	personal communication
$\star m_p$	virophage decay rate	$3.2 * 10^{-1}$	day^{-1}	personal communication
$\star \beta_v$	viral burst size	130	$\frac{\text{viruses}}{host}$	[89]
$\star \beta_{vp} = \rho_{vp}\beta_v$	coinfecting viral burst size	40	$\frac{\text{viruses}}{host}$	personal communication
$\star \beta_p = \rho_p \beta_v$	virophage burst size	1000	$\frac{\text{phage}}{host}$	personal communication
$\star \phi_v$	endocytosis of V rate	$2.2 * 10^{-6}$	$\frac{\text{ml}}{\text{virus}*\text{day}}$	derived; see B
ϕ_p	endocytosis of P rate	$1.1 * 10^{-5}$	$\frac{\text{ml}}{\text{phage}*\text{day}}$	derived; see B
ρ	fraction of infected host offspring	0.5	-	personal communication

spectively. Some parameters shared between the models (denoted with a $\star$) have different values because different reference sets of organisms are used for each model. First principle derivations are shown in B. Overall, these reference parameters are not well constrained based on the current literature and, as a result, we take a sampling approach within large ranges centered around our reference values to analyze the dynamics within both models (see below).

3.2.3 Computational methods

We utilized Latin Hypercube sampling to explore the range of dynamics possible in our models [92]. The sampling ranges were centered about the reference parameter sets in Tables 3.1 and 3.2 and spanned one order of magnitude above and below those values. The only exceptions are ρ_{vp}, ρ_i and ρ, which are bounded between 0 and 1. From these ranges we sample 10^5 points using the midpoints of the hypercubes. We utilized a uniform probability distribution for parameters bounded between 0 and 1 and a log-uniform probability distribution for the other parameters for constructing our hypercubes. We required

Table 3.2: PEM reference parameters. The $\star$ denotes parameters shared between models. The values of the shared parameters may differ between models as they refer to different sets of organisms. When a value is stated for a parameter in the literature for only one system that value is used as the reference for both models as long as it is reasonable. The parameters with "=" show the relation between the parameters present in the model and the free parameters used for sampling. Note "N/A" denotes where information was not available. Hence, in our statistical analysis, we sample from the full range values for ρ_i that retain a reduction in the total burst size of the virus. Personal communication with M. Fischer (MPI-Heidelberg).

Symbol	Meaning	Value	units	Reference
$\star K$	carrying capacity	4.0×10^6	$\frac{host}{ml}$	[90]
$\star b$	host birth rate	1.4	day^{-1}	[90]
$\star d$	host death rate	0.70	day^{-1}	assumed; see B
$\star m_v$	viral decay rate	$3.2 * 10^{-2}$	day^{-1}	[91]
$\star m_p$	virophage decay rate	$3.2 * 10^{-1}$	day^{-1}	personal communication
$\star \beta_v$	viral burst size	300	$\frac{\text{viral particles}}{host}$	personal communication
$\star \beta_{vp} = \rho_{vp}\beta_v$	coinfecting viral burst size	100	$\frac{\text{viral particles}}{host}$	[31]
$\star \beta_p = \rho_p \beta_v$	virophage burst size	1500	$\frac{\text{viral particles}}{host}$	approximated; see B
$\star \phi_v$	absorption of V rate	$4.3 * 10^{-6}$	$\frac{\text{ml}}{\text{viral particles} * \text{day}}$	derived; see B
ϕ_{vp}	rate P attaches to V	$2.2 * 10^{-6}$	$\frac{\text{ml}}{\text{viral particles} * \text{day}}$	derived; see B
$\beta_i = \rho_i \beta_v$	composite burst size	$[0, 1] * \beta_V$	$\frac{\text{viral particles}}{host}$	N/A

$\rho_i + \rho_{vp} \leq 1$ in the PEM to ensure that fewer viruses will be produced during virophage coinfection than without coinfection. We sampled with respect to this constraint by uniformly, randomly sampling between 0 and 1 for each parameter and choosing combinations that satisfied the inequality. We repeated the overall sampling procedure 10 times to give a total of 10^6 sampled points for each model.

For each parameter set, equilibria of the IEM and the PEM were found using Mathematica [93]; script available as Supplementary File 1. Linear stability analysis of coexistence equilibria (all state variables are positive) were also computed in Mathematica. Linear stability analysis of boundary equilibria ($H > 0$, $V > 0$, other state variables are zero) were computed in MATLAB [94]. Similarly, simulations of the models were run using the numerical solvers ode45 or ode15s. All simulations are available as part of Supplementary File 1 and on http://ecotheory.biology.gatech.edu/downloads.

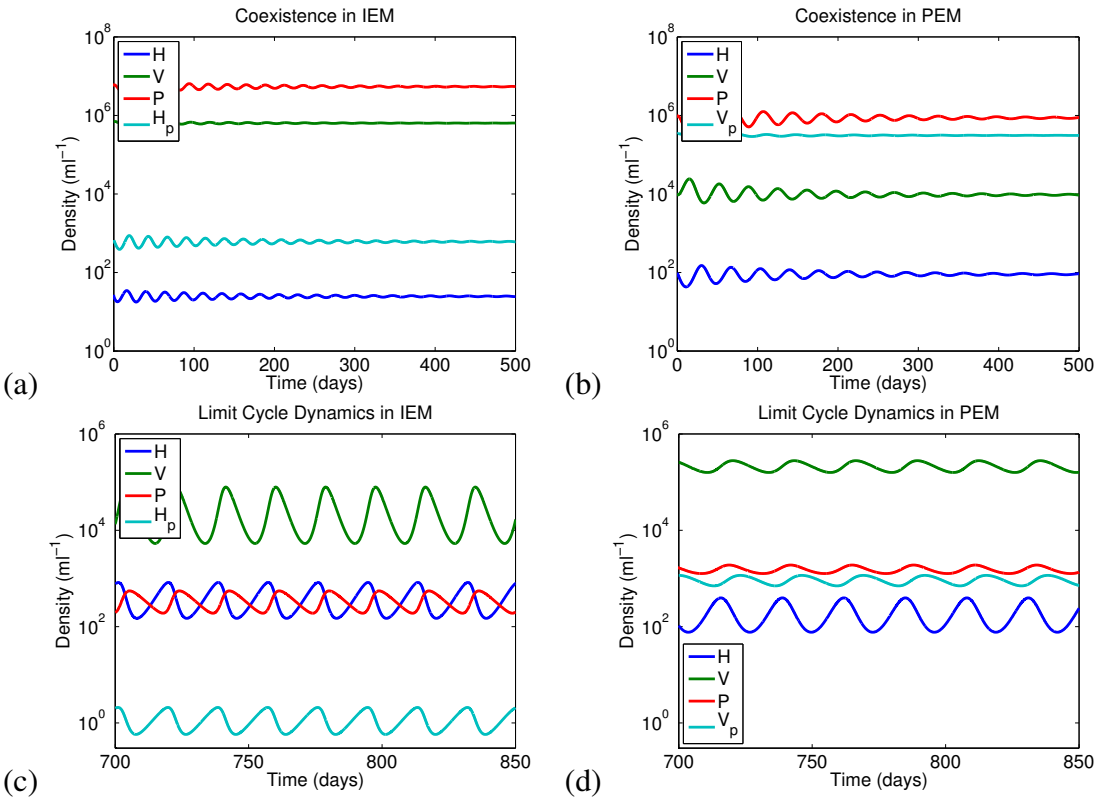


Figure 3.2: Observed dynamics present in both models. The parameter values for each figure are given in B. The initial conditions are small random perturbations from the co-existence equilibria. (a) IEM stable coexistence (b) PEM stable coexistence.(c) Cyclic coexistence in the IEM. (d) Cyclic coexistence in the PEM.

3.3 Results

3.3.1 Stable and cyclical coexistence occur given either biophysical modes of infection

Coexistence equilibria arise in systems (3.1) and (3.2), when the right hand sides of those systems are zero for positive densities of the host, virophage and virus. In both systems, coexistence equilibria can be stable or unstable. We observe for some cases where the coexistence equilibria are unstable that the species exhibit cyclic dynamics. Examples of stable coexistence for both models are shown in Figures 3.2(a,b). Examples of cycle coexistence are shown in Figures 3.2(c,d).

Next we considered the statistical nature of when stable coexistence occurs. We only considered coexistence points where virophage and infected class abundances were each

greater than 10^{-7}ml^{-1}. Boxplots of the parameter distributions in Figure 3.3 show ranges for each parameter value that allow for stable coexistence. For the parameters sampled in log-space the box plots represent the base 10 logarithm of the marginal distributions in terms of the distance from the reference parameter set. For example, consider the box plot for the marginal distribution of the birth rate of the host, b, in the PEM, shown in red (Figure 3.3a). The median of the distribution is nearly 0.5, meaning almost half of the sampled parameter sets for which coexistence occurs have a birth rate over half an order of magnitude ($\sim$3 times) larger than the reference value. Further, the 25th and 75th percentiles of the distributions (edges of boxes in Figure 3a) are above zero, implying that 75% of the sampled parameter sets for which coexistence occurs have a birth rate that is greater than the reference value. For the linearly sampled parameter values (Figure 3.3b), the box plots represent the marginal distributions of the parameters. The reference values for each parameter are marked with an asterisk (*). Overall, the reference values are contained within the middle 50th percentile for most of the parameters; exceptions include b and m_p in the PEM. Coexistence tends to occur when parameters are beneficial to the host and virophage (e.g., high b, high ϕ_p/ϕ_{vp}, high ρ_p, low d, low m_p) and parameters specific to viruses are detrimental (e.g., low ϕ_v, high m_v), when compared to baseline parameter values.

3.3.2 Virophage presence increases host abundance and decreases viral abundance

Histograms of population densities for stable equilibrium points are shown in Figure 3.4. Virophage tend to be the most abundant entity for both models. Additionally, the infected classes (solid cyan lines), which represent the virophage associated with host or virus, tend to be larger than the respective uninfected class for both models (blue line in Figure 3.4a, green line in Figure 3.4b). These results suggest that virophage will be the most abundant entity in field measurement data. However, counterexamples exist where viruses are more abundant than virophage at equilibrium. As a result, field measurements may

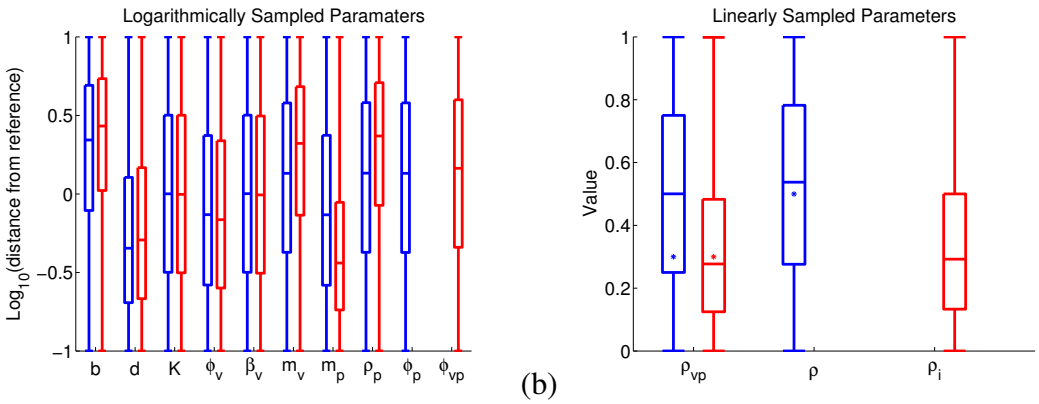
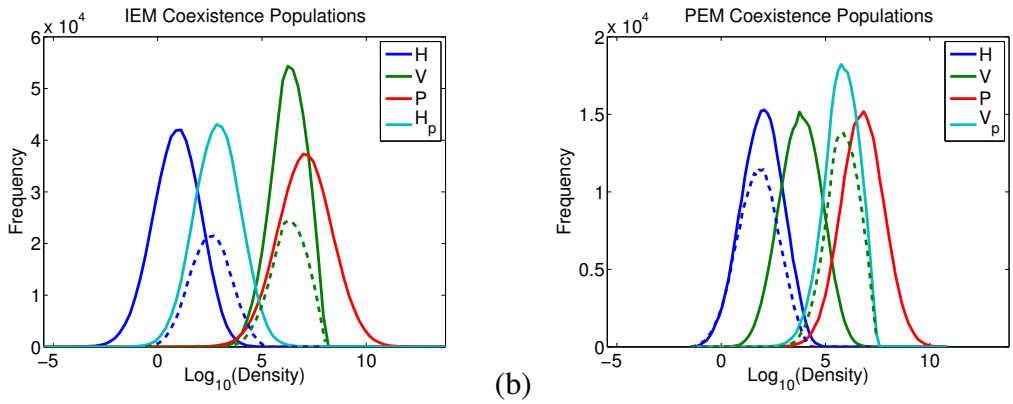


Figure 3.3: Marginal distributions of the parameters for the cases when stable coexistence occurs. Blue (red respectively) boxplots correspond to the IEM (PEM respectively). The median of the distributions are the center lines with the edges corresponding to the 25th and 75th percentile, and the tails extend to the minimum and maximum of the distributions. Shared parameters feature two box-plots for each label and unique parameters feature one box plot. (a) log base 10 distribution relative to the reference parameter sets in Tables 3.2 and 3.1 for logarithmically sampled parameters. (b) Distributions of linearly sampled parameter values. The reference parameter values are denoted by an asterisk (*). No reference parameter value is used for ρ_i.

be useful in determining the covariation between parameters. For example, abundance data could suggest elimination of parameter sets that feature incorrect rank abundance of the populations. The dashed histograms in Figure 3.4 are the population densities for the boundary equilibria. The presence of the virophage causes the host and virus histograms to shift; however, the effect on the total amount of hosts and viruses is not clear and is addressed below.

To address the effect of virophages on the hosts we define the total host abundance as the abundance of host genomes present. For the IEM, a member of the infected class contains a host genome and thus we define the total host abundance as $H^*_{\text{total}} = H^* + H^*_p$, where the (*) denotes equilibrium densities. For the PEM, the hosts are the only modeled variable that involves a host genome and, hence, $H^*_{\text{total}} = H^*$. Similarly, to address the effect of virophages on the viruses we define the total viral abundance as the abundance of viral genomes present. For the PEM, a member of the infected class contains a viral genome and we define the total viral abundance as $V^*_{\text{total}} = V^* + V^*_p$. For the IEM, the



(a) (b)

Figure 3.4: Histograms of stable coexistence populations in each model. Dashed lines are the histograms for the respective boundary equilibrium with host and virus alone. Units for the transformed densities on the x-axis are ml^{-1}. (a) IEM (b) PEM.

viruses are the only modeled variable that involves a viral genome and, hence, $V^*_{\text{total}} = V^*$.

We compare the equilibrium total abundances of the host and viral populations in the presence and absence of the virophage in Figure 3.5. Note we only consider parameter sets where the coexistence equilibrium points are stable. In Figure 3.5, the red lines are the 1-1 line, where virophage has no effect on the host and virus abundances. Since all points in figure 3.5(a,c) lie above the 1-1 line, the virophage increases the equilibrium density of the host. Since all points in figure 3.5(b,d) lie below the 1-1 line, the virophage reduces the abundance of the virus. We note that the relative increases in host abundance and relative decreases in virus abundance tend to be greater in magnitude for the IEM. In total, irrespective of infection mode, the effect of the virophage on equilibrium density can be summarized as the virus of a host's virus is the host's "friend."

To see why virophage always reduce the total viral density and increase the total host density at equilibrium, consider the respective dynamics. We define an average burst size parameter: $\bar{\beta}_v = \frac{\beta_v H + \beta_{vp} H_p}{H + H_p}$ for the IEM, and $\bar{\beta}_v = \frac{\beta_v V + (\beta_{vp} + \beta_i) V_p}{V + V_p}$ for the PEM. Note that $\bar{\beta}_v \leq \beta_v$ in both models. The dynamics for the total host and total virus densities simplify to a standard competitive Lotka-Volterra system where the virus is a predator of the host (derived in B.0.2):

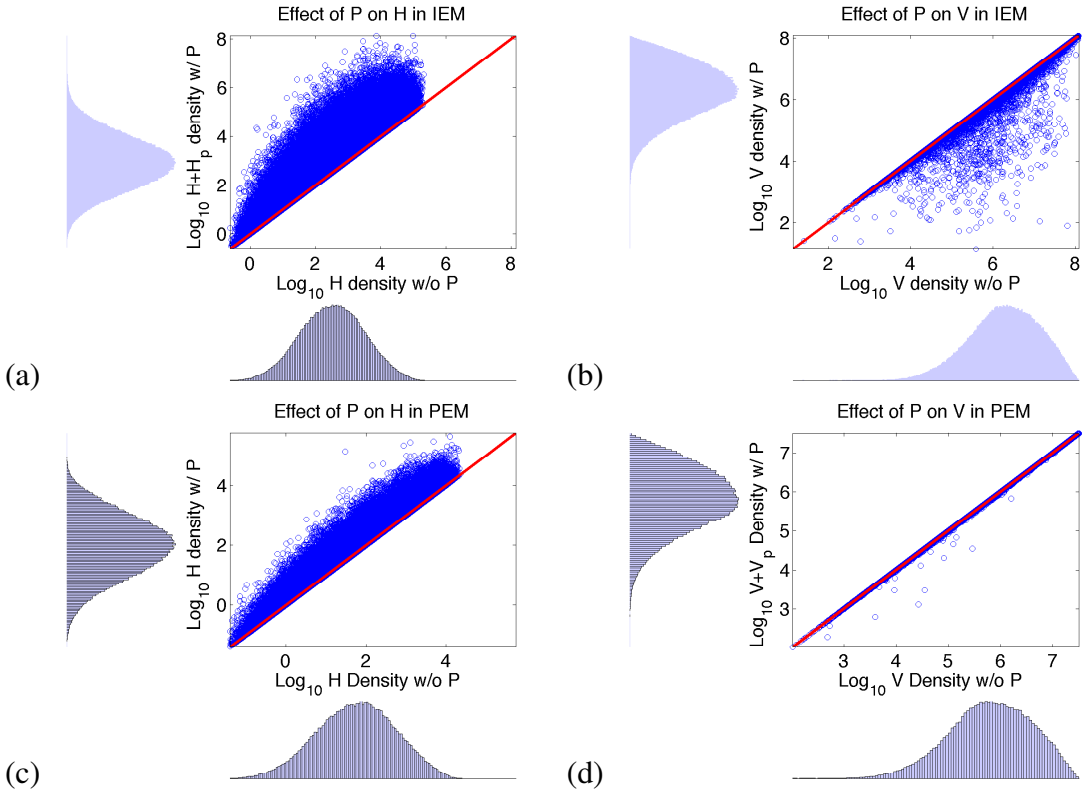


Figure 3.5: Comparing the marginal distributions of the species' genome abundances at the coexistence equilibria and the boundary equilibria where only the host and virus are present. Infected classes are combined with uninfected classes for comparison of total genome abundances (e.g., $V_{\text{total}} = V + V_p$ in the PEM). Units for the transformed densities on each axis are ml^{-1}. Effect of virophage on (a) host genome abundance and (b) viral genome abundance in the IEM. Effect of virophage on (c) host genome abundance and (d) viral genome abundance in the PEM.

$$\begin{aligned}\dot{H}_{total} &= H_{total}\left[b - d\left(1 + \frac{H_{total}}{K}\right)\right] - \phi_v V_{total} H_{total} \\ \dot{V}_{total} &= \bar{\beta}_v \phi_v H_{total} V_{total} - m_v V_{total}.\end{aligned} \tag{3.3}$$

The parameters of system (3.3) are the same as in models (3.1) and (3.2). Solving for the equilibrium populations (denoted with $*$) and utilizing the bound on $\bar{\beta}_v$, we obtain:

$$H^*_{total} = \frac{m_v}{\phi_v \bar{\beta}^*_v} \geq \frac{m_v}{\phi_v \beta_v} = H^*_b \tag{3.4}$$

$$\begin{aligned} V^*_{total} &= \frac{1}{\phi_v}\left[b - d\left(1 + \frac{m_v}{\phi_v \bar{\beta}^*_v K}\right)\right] \\ &\leq \frac{1}{\phi_v}\left[b - d\left(1 + \frac{m_v}{\phi_v \beta_v K}\right)\right] = V^*_b, \end{aligned} \tag{3.5}$$

where the subscript b refers to the boundary equilibrium with hosts and viruses only. The burst size of the genome level model at equilibrium is represented by $\bar{\beta}^*_v$. Thus, virophage coinfection causes a reduction in burst size, which in turn, increases total host abundance and decreases total viral abundance. This effect occurs so long as the virophage has a deleterious effect on the burst size of the virus, i.e., $\beta_{vp} < \beta_v$ for the IEM and $\beta_{vp} + \beta_i < \beta_v$ for the PEM (see B.0.2).

3.3.3 Bistability in PEM

In our numerical simulations we did not find parameter values for which bistability arises in the IEM. In contrast, approximately 1% of the parameter sets in the PEM yielded bistability. Bistability arises when two equilibria are locally stable, and results in asymptotic dynamics dependent on initial conditions. An example of this bistability between two ranges of time is shown in Figure 3.6a. When smaller amounts of virophage are added, the virophage are unable to invade (two dashed curves), whereas when amounts greater than a certain

amount (here $\sim 10^{4.5}\text{ml}^{-1}$) are added the virophage are able to invade (three solid curves). In most cases where bistability was observed, the boundary and the coexistence equilibria were both locally stable fixed points. In these cases, there also existed a second coexistence equilibrium that was saddle point (i.e., semistable). In a few other cases, bistability arose when the boundary equilibrium was stable and both of the coexistence equilibria were locally unstable. In these few cases, there was cyclic coexistence between the host, virophage and virus. We note that this case was difficult to find numerically. Hence, in the following we focus on parameter sets where one coexistence equilibrium and the boundary equilibrium are locally stable.

For parameter sets with bistability there exists a basin of attraction for the coexistence point in phase space. We interpret the size of this basin of attraction as a proxy for the robustness of coexistence to environmental perturbations. Hence, we identified the boundary of this basin of attraction along the axes of phase space from the boundary equilibrium (Figure 3.6b) and from the coexistence point (Figure 3.6c). The boundary of the basin of attraction has a different interpretation in each case, as discussed individually below.

Figure 3.6b shows a histogram of the boundary of the basin of attraction along the virophage axis from the boundary equilibrium. An interpretation of this boundary in phase space is the minimum amount of virophage required to invade a system at equilibrium with hosts and viruses alone. These results were obtained by randomly sampling 1000 parameter sets where bistability is expected to occur based on the linear stability analysis. We repeatedly simulated the dynamics with the boundary equilibrium for hosts and viruses and varying amounts of virophage as the initial condition. We performed a bisection method in log space for initial amounts of virophage with a range of $[10^{-4}, 10^{4}]$ times the virophage population at the coexistence equilibrium. Out of the 1000 samples, for 93 simulations either the basin of attraction was outside the range of our bisection method or the dynamics did not converge to virophage invasion or crashing within a specified time. We did not include these parameter sets in our histogram giving a total of 907 parameter sets in the data.

For the remaining parameter sets, the average of the minimal amount of virophage added that led to coexistence and maximum amount of added virophage that led to virophage extinction are the values in the histogram. These values are accurate within .005 in the log space range of the prefactor as mentioned above. Overall, this figure illustrates that a non-negligible amount of virophage must be introduced in order for coexistence to occur.

Figure 3.6c shows histograms of the boundary of the basin of attraction along the phase space axes from the coexistence equilibrium for our 1000 samples. The values of the x-axis are relative to the respective coexistence equilibrium population. An interpretation of this boundary in phase space is a bound on the amount of each respective population that can be added or removed without causing the virophage to crash. These values were obtained using a bisection method similar to the one previously described. One difference is a smaller range was used ($[10^{-1}, 10^{1}]$ times the respective coexistence population). This range spans from reducing the respective population to 10% of its coexistence value to increasing the respective population to 10 times its coexistence value. We bin together the parameter sets for which the boundary existed outside our range and include this in our histograms.

Figure 3.6c shows that coexistence is differentially robust to perturbations of the different population densities. For example, coexistence is very robust to removal of virophage as nearly all of the parameter sets maintained coexistence within the entire range of perturbations. In comparison, the peak of the V_p histograms are close to 0 and almost entirely contained within our range. This suggests virophage coexistence, when it occurs ina region of parameter space corresponding to bistability, is highly sensitive to addition or removal of virus-virophage composites.

3.3.4 Sampled coexistence points tend to be attracting

We now investigate the frequency of stable coexistence in our parameter sets. We used linear (local) stability of the coexistence equilibria and the boundary equilibria (equilibria

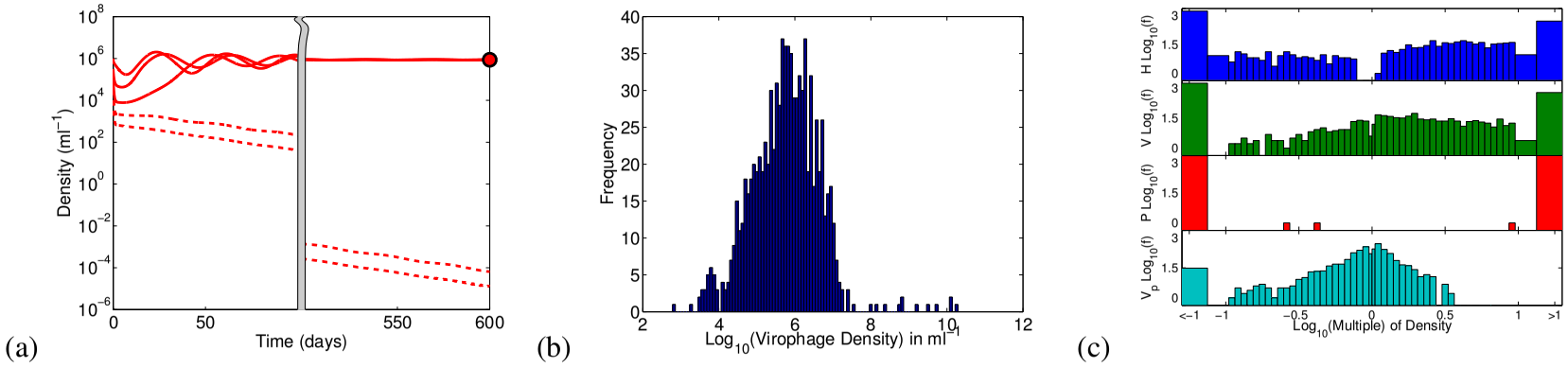


Figure 3.6: Bistability in the PEM. (a) Virophage dynamics given different initial conditions. The initial conditions are that the host and virus are at the boundary equilibrium ($H_0 \approx 23.0$ and $V_0 \approx 3.20 * 10^5$) and the virophage varies as $P_0 = 10^m$ where $m \in \{4, 4.5, 5, 5.5, 6\}$. Dynamics that lead to coexistence (virophage crashing) are plotted with solid (dashed) lines. The red circle shows the theoretical virophage coexistence population. (b) Histogram of the approximate distance from the boundary equilibrium to the edge of its basin of attraction in the direction of virophage density. We present the data for 907 random parameter sets that yielded convergence. (c) Histograms of the logarithm of frequencies of approximate locations of the basin of attraction boundaries of the coexistence equilibrium with respect to each population. Histograms of the approximate distance from the coexistence equilibria to the edge of its basin of attraction in the direction of the different state variables. Data represents 1000 random parameter sets that allow bistability and where the basin of attraction boundaries lie between 0.1 and 10 times the respective coexistence population densities. Distances were computed in terms of a multiple of the respective coexistence equilibrium population density. The abscissa is the logarithm of the multiple of the respective coexistence equilibrium population density. Values below (above) 0 mean that the boundary is located at density values less (greater) than the coexistence equilibrium density. The ordinate is the logarithm of the number of parameter sets. Points that lie outside the range of 0.1 and 10 times the equilibrium density are binned together and denoted as < -1 and > 1 respectively.

with host and virus alone) to identify if the equilibria were stable ("S") or unstable ("U"). A statistical enumeration of the linear stability of the equilibria is shown in Table 3.3. Table 3.3a corresponds to the IEM. The columns of Table 3.3a define the stability of the boundary equilibrium of the IEM and the rows of Table 3.3a define the stability of the coexistence point of the IEM. Table 3.3b corresponds to the PEM. The columns of Table 3.3b define the stability of the boundary equilibrium of the PEM. Since multiple coexistence equilibria can arise in the PEM, the rows of Table 3.3b are divided into cases where there is one coexistence equilibrium ("single coexist") or two coexistence equilibria ("multi coexist").

We make a few points about Table 3.3. First, stable coexistence occurs when at least one of the coexistence equilibria is stable. When all of the coexistence equilibria are unstable, then cyclic coexistence, aperiodic coexistence, or extinction of the virophage are possible outcomes. Second, out of the 10^6 parameter sets, coexistence equilibria (either stable or unstable) are observed for approximately half of the parameter sets for the IEM and approximately a quarter of the parameter sets for the PEM. Thus, based on these totals, a larger portion of the parameter space allows for coexistence in the IEM versus the PEM.

Third, since the "SS" row in Table 3.3b sums to zero, bistability between coexistence points was not observed in the PEM. Thus, for our parameter ranges, bistability only occurs when both a boundary equilibrium and a coexistence equilibrium are locally stable. Since bistability was not observed in the IEM, we interpret these results to suggest that the coexistence of virophage is more robust to perturbations in the IEM versus the PEM. Finally, cyclic coexistence can only occur when all coexistence equilibria are unstable. Stable coexistence equilibria occur at a higher frequency than unstable equilibria in the IEM (compare "S" row and "U" row in Table 3.3a). Similarly, in the PEM, it is more frequent that at least one coexistence equilibrium is stable than all coexistence equilibria being unstable (compare "S" and "SU/US" rows to "U" and "UU" rows in Table 3.3b). Thus, the dynamics observed in our models suggest that if cycles are observed in experimental population dynamics, then either the parameters of the systems are finely tuned or the cycles are

Table 3.3: Frequency of stable coexistence in the IEM and PEM models. Each table lists the number of parameter points satisfying different combinations of linear stabilities for the coexistence points and the boundary equilibria for (a) IEM and (b) PEM. The total number of sampled parameter sets for each model is 10^6. "S" refers to stable and "U" refers to unstable. (b) In the PEM, parameter sets can have one("Single Coexist") or multiple("Multi Coexist") equilibrium points. Rows with multiple letters denote the stabilities of the two coexistence points.

(a)

IEM		Boundary	
		S	U
Coexist	S	0	498821
	U	0	34
	Total	0	498855

(b)

PEM		Boundary	
		S	U
Single Coexist	S	0	243324
	U	0	9131
Multi Coexist	SS	0	0
	SU/US	11342	0
	UU	77	0
	Total	11419	252455

forced by sources outside of our modeling framework (e.g., predation on the host species).

3.4 Discussion

The study of virophage interactions with viruses is at its infancy. Nonetheless, multiple independent discoveries have been made of virophage populations persisting with viruses and their eukaryotic hosts in a diverse range of environments from cooling towers to the open oceans. These discoveries motivated our central aim to develop biophysically-motivated models of the interactions among virophage, viruses and their eukaryotic hosts and to understand what effects virophage have on the dynamics of these populations. The models we proposed correspond to two distinct cases: (i) where virophage attach to viruses and then the composite infects host cells (the PEM); and (ii) where viruses and virophage independently infect host cells (the IEM). Coexistence amongst all populations is possible in both models when analyzed over plausible ranges of parameter space. In addition, both models allowed for stable and cyclical asymptotic dynamics. Importantly, we demonstrated both analytically and numerically that so long as virophage negatively affect virus burst size then virophage will act as the "friend" of their hosts, i.e., increasing host abundance and decreasing virus abundance, irrespective of infection mechanism.

These results add an ecological layer to prior observations of the cellular level effects of virophage on viruses and eukaryotic hosts. They also suggest testable hypotheses for evaluating the ecological effects of the presence of virophage within communities (e.g., when virophage enter a new environment, they should drive viral populations down resulting in an increase of host populations). The models may also be used to help distinguish between the biophysical mode of infection where virophage are present. For example, bistability was observed only in the PEM. An experimentalist may test for bistability by first obtaining coexistence between virus and host within a chemostat and then observing both extinction and coexistence of virophage after introducing different concentrations of virophage. If bistability is observed with virophage that follow the PEM then stochastic fluctuations may be more likely to lead to virophage extinction in comparison to systems where the virophage follow the IEM.

An alternative approach to distinguishing between infection modes arises from analyzing phase lags in those instances where cyclical dynamics are observed. A similar approach has been proposed to distinguish between indirect and direct transmission in the spread of infectious pathogens within traditional epidemiological SIR-type models [85]. Here, we have limited preliminary evidence to suggest a similar approach may also be of use. We observed that the virus population cycles preceded the virophage population cycles in the IEM, whereas the virophage cycles preceded the virus cycles in the PEM. By "precede", we mean that the population maximum (and minimum) of one type appears immediately before the population maximum (and minimum) of the other type. Hence, measurements the densities and orderings of peaks for the virus and virophage could help distinguish between infection modes. However, our analysis involves a small number of examples; initial exploration is presented in B and warrants follow-up study.

Although the models developed here were constructed with virophage in mind, they may be useful in modeling the ecological effect of satellite viruses or other defective interfering particles. In fact, a previous model of defective interfering particles shares a

similar form to our independent entry model; however, it differs in construction and analysis whereby infections were treated as stage structured and simulations assumed an *in vitro* setting where viruses were repeatedly introduced through passages [95].

These models may also be relevant to other organisms since virophage function as a special case of hyperparasitism. In our case the virophage functions as the hyperparasitoid. Previous models of hyperparasitism have been limited to two classes of models: epidemiological type models where the parasite and hyperparasite spread through direct transmission of the hosts [96, 97] and population models based on difference equations [98, 99].

In moving forward, it is important to note a secondary contribution of this study: the establishment of parameter baselines applicable to distinct biophysical modes of virophage-virus-host interactions. In sampling parameter space we assumed parameter distributions independent from each other. In reality, both the range of parameter values and their covariation are likely to be more constrained as a result of trade-offs, biophysical limits, and other effects. We suggest the need for further empirical studies to refine both the qualitative and quantitative nature of these interactions. Such refinement is likely to provide further evidence to establish when environments are likely to support a virophage population in the first place, identify ecological effects common to both modes, and identify which of our proposed means for distinguishing the mode of coinfection from population level data are useful. Additionally, given better constraints on the potential range of parameter values, we will also be able to extend the current model to ask evolutionary questions, e.g., how virophage interactions may evolve in distinct ecological contexts. In doing so, extending the current framework to a spatially explicit context is likely to be of use as spatial models stabilize viral-host systems and can yield alternative conclusions to evolutionary questions [42]. Extension to a spatial model seems particularly relevant for virophage given the requirement of host coinfection within a large, complex population (whether infecting together or independently) in order for virophage to reproduce.

CHAPTER 4
VIROPHAGE FEASIBILITY IN SPATIAL MODELS

Broader Context: This chapter connects the prior two chapters. By considering spatial dynamics we show that virophage feasibility is less likely for high viral adsorption rates when the virophage associates with the virus first. We demonstrate this by calculating the effect of population structure on fitness. We introduce a novel computational methodology to calculate this spatial dependence. This method could be applied in other ecological settings. By connecting the estimates of fitness to the local population structure, we initiate considering dynamics from the basis of pair-correlation profiles. This is complementary to common theoretical ecology techniques of recasting spatial dynamics in terms of the evolution of the one-point ("mean") and two-point ("pair") densities. Overall, this chapter demonstrates how the effect of space can depend on the order of interactions in a tri-partite network.

4.1 Introduction

Hyperparasitic viruses, virophage, must coinfect a eukaryotic host along with a giant virus in order to propagate. It is hypothesized that coinfection intitiates through one of two associative mechanisms: the paired entry mode (PEM) or the independent entry mode (IEM) [31, 28, 100]. In PEM the virophage associates with its viral partner first and the pair later coinfect the mutual host. In IEM the virophage enters the host with coinfection proceeding following infection by the viral partner. Prior theoretical work has shown that in well-mixed conditions both modes of coinfection are ecologically stable albeit with different robustness [100]. Here we ask how does spatial structure affect the feasibility of both modes of coinfection?

Spatial structure can qualitatively and quantitatively affect the outcome of ecological dynamics as compared to well-mixed models [36]. For example, the population-level oscillations between predator and prey dampen as a result of decoherence in oscillation across space [101]. Also, space can facilitate coexistence in predator-prey, host viral systems [42]. The mechanism for increased coexistence was the existence of spatial refuges: density fluctuations across space allowed increased host growth rates in low virus regions. Extending these results to virophage is non-trivial as virophage should rely on the local presence of both hosts and viruses. However, an understanding of hyper-parasitic dynamics in space is largely unexplored. To our knowledge only one previous work addressed the role of pattern formation on a hyper-parasitic system [102]. The model is not applicable to the virophage system because it focused on a system where the mutual host is a sessile plant. Additionally, we are interested in the effects of spatial dynamics on feasibility of the hyper-parasite–not just the conditions for pattern formation.

We address the role of space on virophage coexistence by simulating spatial individual based models across ranges of biologically reasonable parameter space. In doing so we consider two distinct virophage modes of coinfection: IEM and PEM. We compare the feasibility results from the spatial model to expectations from an ODE well-mixed model. We finish by investigating the role of population composition in space on the local fitness of virophage.

4.2 Methods

4.2.1 Spatial model

We develop two stochastic, spatial IBMs corresponding to virophage dynamics by either IEM or PEM. The domain is a two-dimensional, periodic 100 by 100 square lattice where multiple hosts, viruses, and virophage can occupy lattice points. Dynamics occur at fixed time steps given stochastic processes that include diffusion, host growth, host death, viral infection of hosts, virophage association, infected host lysis, and virophage and viral de-

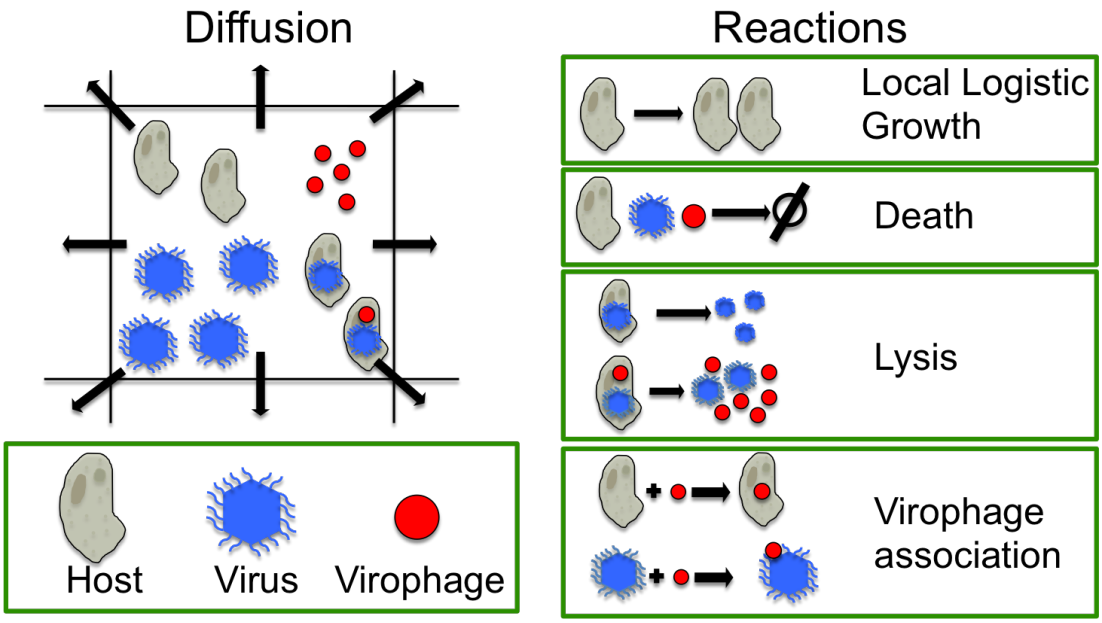


Figure 4.1: Virophage Reaction-diffusion dynamics modeled on a lattice. Dynamics are stochastic and can occur at fixed time steps. Only one virophage association reaction occurs in each model depending on whether the virophage associates with the host or the virus

cay (see Figure 4.1). Note, complex host movement is appropriately modeled as simple diffusion for the length scale considered for lattice points [103]. Each process is Poisson such that the probability of an event occurring during a time-step follows an exponential distribution. Simulation parameters are shown in Table 4.1

We initiate the spatial dynamics by randomly distributing hosts, viruses, virophage given that the total initial abundances match numerical equilibrium solutions of the analogous mean field model. If the long-term dynamics are oscillatory, e.g., a limit cycle, the initial condition is a time average of the populations over the orbit. Each simulation is run for 9.6×10^4 timesteps corresponding to 200 days given our simulation parameters. Dynamics were inspected to ensure that transients had been eliminated.

4.2.2 Mean Field ODE models of coinfection

Here we present mean-field ODE models for PEM and IEM modes of coinfection. The dynamics are the same as in the spatial model with the absence of diffusion. These models

Table 4.1: Parameter values for both IEM and PEM models. The parameters are the same in the spatial models and the ODE models. The adsorption rates are specified throughout the chapter.

Parameter	Value	Units	Interpretation
r	0.0625	hr^{-1}	host growth rate
K	$4.0 * 10^6$	ml^{-1}	carrying capacity
K/lattice point	32	hosts	carrying capacity per lattice point
m	$\frac{r}{8} = 0.0078$	hr^{-1}	decay/death rate
m_p	$\frac{10r}{8} = 0.0781$	hr^{-1}	virophage decay rate
λ	0.0417	hr^{-1}	lysis rate
β	20	viruses	viral burst size
β_{vp}	6	viruses	viral burst size (coinfection)
β_p	140	virophage	virophage burst size
D_p	2.04E-4	$\text{cm}^2\ \text{hr}^{-1}$	virophage diffusion
D_v	2.86E-5	$\text{cm}^2\ \text{hr}^{-1}$	viral diffusion
D_h	4E-6	$\text{cm}^2\ \text{hr}^{-1}$	host diffusion (amoeba)
τ	$\frac{1}{20}$	hr	time step
Ω	100×100	lattice points	domain size
Δx	2E-2	cm	lattice point length

were used to determine the initial conditions and expected feasibility of virophage. The PEM dynamics are:

$$
\begin{aligned}
\dot{H} &= \overbrace{rH\left(1 - \frac{H+I+C}{K}\right) - mH}^{\text{logistic growth}} - \overbrace{\phi(V+V_p)H}^{\text{viral infection}} \\
\dot{I} &= \overbrace{\phi V H}^{\text{viral infection}} - \overbrace{\phi V_p I}^{\text{associated virus infection}} - \overbrace{(d+\lambda)I}^{\text{death \& lysis}} \\
\dot{V} &= \overbrace{\lambda(\beta I + \beta_{vp} C)}^{\text{infected and coinfected host lysis}} - \overbrace{\phi V\,(H+I+C)}^{\text{viral infection}} - \overbrace{\phi_p V P}^{\text{virophage association}} - \overbrace{mV}^{\text{decay}} - \overbrace{m_p V_p}^{\text{virophage decay}} \\
\dot{P} &= \overbrace{\beta_p \lambda C}^{\text{coinfected host lysis}} - \overbrace{\phi_p(V+V_p)P}^{\text{virophage association}} - \overbrace{m_p P}^{\text{virophage decay}} \\
\dot{C} &= \overbrace{\phi V_p(H+I)}^{\text{associated virus infection}} - \overbrace{(d+\lambda)C}^{\text{death \& lysis}} \\
\dot{V_p} &= \overbrace{\phi_{vp} V P}^{\text{virophage association}} - \overbrace{\phi V_p\,(H+I+C)}^{\text{associated virus infection}} - \overbrace{(m_v + m_p)V_p}^{\text{decay}}
\end{aligned}
\tag{4.1}
$$

where H tracks uninfected hosts, I_j tracks hosts infected by the virus alone, V tracks the viruses, P tracks the virophage, C tracks hosts coinfected by both virus and virophage, and V_p tracks the associated class for the PEM–a virophage associated with a virus. Infected and coinfected hosts act as a sink for viruses. Viruses are tracked either as unassociated, V, or associated with virophage, V_p. Thus associated viruses act as a virophage sink. Additionally, virophage decay can occur even when associated with a virus transitioning associated viruses to unassociated viruses.

The IEM dynamics are:

$$
\begin{aligned}
\dot{H} &= \overbrace{r(H+H_p)\left(1-\frac{H+H_p+I+C}{K}\right)-mH}^{\text{logistic growth}} - \overbrace{\phi V H}^{\text{viral infection}} - \overbrace{\phi_p P H}^{\text{virophage infection}} + \overbrace{m_p H_p}^{\text{virophage decay}} \\
\dot{I} &= \overbrace{\phi V H}^{\text{viral infection}} - \overbrace{\phi_p P I}^{\text{virophage infection}} - \overbrace{(d+\lambda) I}^{\text{death \& lysis}} \\
\dot{V} &= \overbrace{\lambda(\beta I + \beta_{vp} C)}^{\text{infected and coinfected host lysis}} - \overbrace{\phi V\left(H+H_p+I+C\right)}^{\text{viral infection}} - \overbrace{mV}^{\text{decay}} \\
\dot{P} &= \overbrace{\beta_p \lambda C}^{\text{coinfected host lysis}} - \overbrace{\phi_p (H+H_p) P}^{\text{virophage association}} - \overbrace{m_p P}^{\text{virophage decay}} \\
\dot{C} &= \overbrace{\phi V H_p}^{\text{virus infecting associated host}} - \overbrace{(d+\lambda) C}^{\text{death \& lysis}} \\
\dot{H}_p &= - \overbrace{m H_p}^{\text{associated host death}} - \overbrace{\phi V H_p}^{\text{viral infection}} + \overbrace{\phi_p P H}^{\text{virophage infection}} - \overbrace{m_p H_p}^{\text{virophage decay}}
\end{aligned}
\tag{4.2}
$$

where H tracks uninfected hosts, I_j tracks hosts infected by the virus alone, V tracks the viruses, P tracks the virophage, C tracks hosts coinfected by both virus and virophage, and H_p tracks the associated class for the IEM–a virophage associated with a host. Infected and coinfected hosts act as a sink for viruses. Uninfected hosts are tracked either as unassociated, H, or associated with virophage, H_p. Thus associated viruses act as a virophage sink. Additionally, virophage decay can occur even when associated with a host transitioning associated hosts to unassociated hosts. Note, we parametrize both models such hosts decay

at the same rate as host death. This ensures quantitative differences between models result from the mode of coinfection alone.

4.3 Results

4.3.1 Feasibility of parameter space between well-mixed and spatial models

We independently vary the viral adsorption rate, ϕ, and the virophage association rate, ϕ_{vp} or ϕ_p. A parameter set is feasible for the spatial model if virophage, virus, and hosts coexist at the end of simulation. A parameter set is feasible for the ODE model if virophage, virus, and hosts each have a greater density than 1 particle/domain at (dynamic) equilibrium for all populations. Thus if the ODE model leads to sufficiently large oscillations the parameter set is labeled as infeasible. The feasibility of IEM in both spatial and ODE models is shown in the top row of Figure 4.2. Space does not considerably affect the ranges of feasible parameter space for this model. The feasibility of PEM in both spatial and ODE models is shown in the bottom row of Figure 4.2. The feasible parameter space is enriched (reduced) for spatial dynamics for low (high) viral adsorption rates ϕ.

Interestingly, in the spatial coexistence phase-plot there appears to a value of $\phi = 10^{-8.5}$ that is infeasible whereas larger and smaller values of ϕ are feasible. We investigated this further by considering $\phi_p = 10^{-7.4}$ and $\phi_p = 10^{-7.0}$ and performing 40 additional simulations each. For this extra simulations we randomized the values of the initial abundances. The initial abundances were randomly selected from a uniform distribution spanning from 0.9 times the minimum abundance and 1.1 times the maximum abundance observed across the 4-neighborhood of initial abundances in the phase-space. None of the simulations led to coexistence suggesting the observation of infeasibility is robust, although we do not yet understand the mechanism.

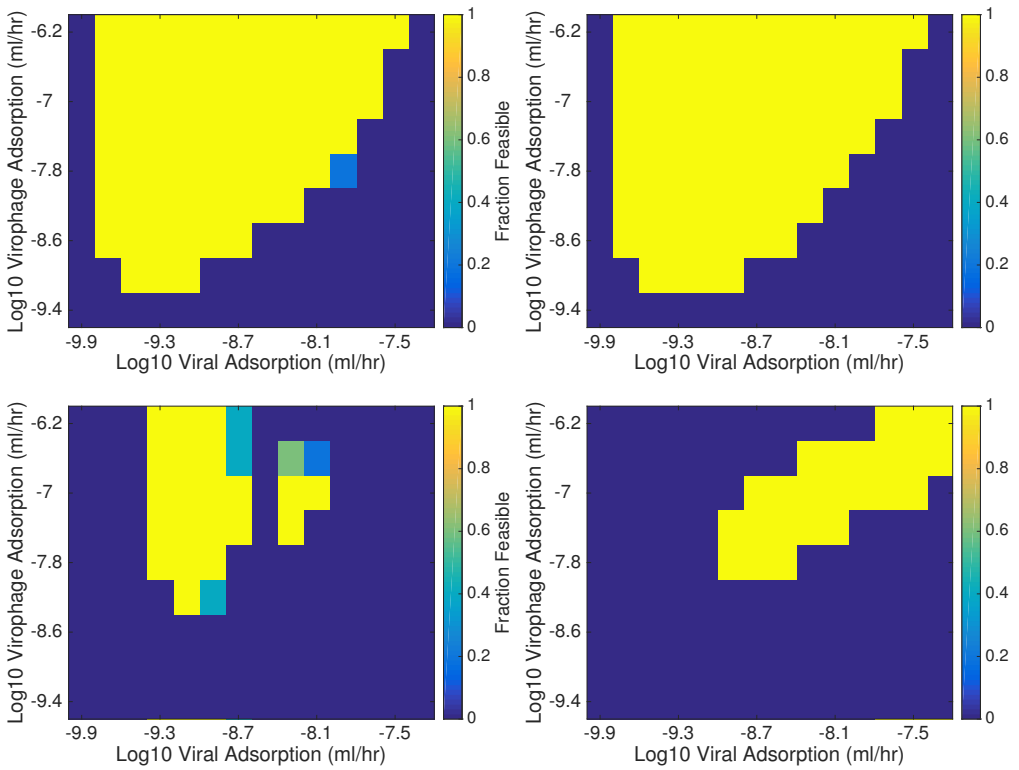


Figure 4.2: Feasible parameter space for virophage across models. The feasibility of virophage dynamics for the (top row) IEM (left) spatial model (Left) and (Right) ODE model. The feasibility of virophage dynamics for the (Bottom row) PEM (Left) spatial model and (Right) ODE model. For the spatial model, the value is the fraction of 5 simulations that led to coexistence of virophage, virus, and host after 96000 time steps. For the spatial model, the value indicates whether the parameter values lead to either stable coexistence or oscillations such that no population is below the value of 1 per domain size.

4.3.2 Differential virophage fitness across space

An individual's fitness depends on it's properties (e.g., genotype/phenotype) and the environment it inhabits. Here, we explore how virophage fitness depends on the nearby composition of the populations. The number of future coinfected lysis events that arises from single coinfected lysis event, or effective reproduction number, R_{eff}, is the one measure of fitness. This is because virophage infections only matter when the host first becomes coinfected. All subsequent virophage infections have no dynamical impact other than acting as a sink. In a well-mixed system if $R_0 > 0$ then the associated population grows. Hence, the spatial dependence of R_{eff} indicates the areas where virophage growth occurs. We track

R_{eff} of the coinfected hosts, as opposed to of the virophage population, because viruses produced during the burst may associate with an external virophage and contribute to the virophage fitness.

We consider parameter sets with maximal viral adsorption rate while maintaining co-existence across our replicates ($\phi = 10^{-8.1}$ ml hr^{-1} and $\phi_{vp} = 10^{-7}$ ml hr^{-1}). Dynamics for this parameter set are shown in Figure 4.3. The populations oscillate in time. Coupled with oscillations in abundance are spatial oscillations leading to a dynamic checkerboard pattern (see Figure 4.3 for snapshots of virus dynamics at lowest and highest abundances). We choose the final state of our simulations as the initial condition for our analysis of the spatial dependence of R_{eff}.

Pair correlation profiles quantify the abundance of a population at specified distances away from focal lattice points relative to the total abundance of the population [104]. The pair cross-correlation profiles between coinfected hosts and all other populations are shown in Figure 4.4. The value of the pair correlation is relative such that a value greater (less) than 1 indicates larger (smaller) abundances at some distance away as compared to the expectation due to random distribution of populations. Coinfected hosts are near regions of high abundance of populations associated with virophage (C, V_p, and P). Coinfected hosts are also near regions of high abundance of host populations and low abundances of viral populations. This is a consequence of local enhancement of fluctuations arising due to viral bursts. It is analogous to prior results in the spatial model of virus-host dynamics without virophage [100].

We computationally assess the spatial variation of virophage fitness. We individually induce lysis across 100 randomly sampled lattice points that contain coinfected hosts. As the spatial dynamics run forward, we track both viral and virophage progeny from this bursting event. We continue tracking each viral particle as long as it does not leave the system through a sink. For example, we would track a progeny virus as it associates with a virophage (regardless of origin) and then continue tracking as it infects either an infected

or uninfected host. We would not continue tracking the virus had it infected a coinfected host. The total number of secondary lysing coinfected hosts is reported as our measure of fitness, R_{eff}. We repeat the simulation 50 times to calculate an ensemble average of the fitness for each induced lysis event.

For each focal lattice point, we also repeated the forward simulation procedure but while resorting all other lattice points. This allows identifying the relative importance of the population composition of the individual lattice point versus the population composition of local neighborhood. Figure 4.5 shows the distribution of R_{eff} values between the standard initial space and the randomized space. The figure is split into quadrants that demarcate the relative importance of the composition of the focal lattice point versus the neighborhood to the fitness. For example, the top left quadrant indicates points where the local neighborhood increases fitness despite a composition of populations that reduce fitness in the focal lattice point. The smaller range of R_{eff} values across the axis corresponding to randomizing the entire lattice suggests that the local neighborhood plays a major role in determining the fitness of virophage.

To address the impact of neighborhood composition on fitness, we focus on the lattice points with large differences between fitnesses in the full space and the randomized spaces. First, we consider points where the local environment increases fitness despite the population composition of the focal lattice point and where the local environment decreased fitness despite the population composition of the focal lattice point. This corresponds to points in the top right and the bottom left quadrant respectively in Figure 4.5. The pair cross-correlation profiles for these focal lattice points are shown in Figure 4.6. Proximity to hosts and uninfected hosts increases fitness. Proximity to other coinfected cells does not have a considerable effect. Proximity to viruses increases fitness where as proximity to viruses associated with virophage decrease fitness.

To gain further intuition, we also show the pair cross-correlation profiles for the lattice point with the largest $R_0 = 3.875$ in Figure 4.6. The population composition of its

neighborhood largely matches the curve corresponding to neighborhood increasing R_{eff}. Contrary to the trends the maximum R_{eff} lattice point is proximate to a very low abundance of virophage and viruses associated with virophage. This suggests it is beneficial when a coinfected host has randomly dispersed into an area recently unaffected by virophage and thus has a large unassociated viral population.

We report one more observationThe initial state of the system, with blue (red) points where local neighborhood increases (decreases) fitness, is shown in Figure 4.7. Points where the neighborhood either increases or decreases fitnesses tend to be located on the periphery of areas lacking some population type.

4.4 Discussion

Virophage engage in a unique hyper-parasitic relationship in a tri-partite system. In order to propagate virophage must coinfect a host with another virus. This effectively involves a tri-molecular reaction, which is complicated in space as populations correlate and anti-correlate with each other. We addressed the impact of spatial dynamics of virophage coexistence. We showed that feasibility of virophage that associate with hosts is comparable in the spatial system to the well-mixed system. Meanwhile, feasibility of virophage that associate with viruses increases for low viral adsorption rates and decreases for high adsorption rates. We attribute the decrease in feasibility due to the emergence of spatial clustering and the long term instability of these cluster dynamics.

In order to evaluate the effect of clustering on fitness, we introduced a computational method to evaluate the spatial variation in R_{eff}. We are not aware of prior efforts to quantify spatial variation of fitness in virus-microbe systems (although see [105] for an abstracted, general treatment). While our focus here was on exploring the robustness of virophage coexistence in light of clustering, our introduced methodology has wider applications. For example, our fitness differences arose purely to the population composition within neighborhoods. However, fitness differences arising from mutations may also be

considered. Our method provides a framework to address how ecology through demography can impact evolution. For example, our method may be particularly relevant to studying the emergence of antibiotic resistance given the importance of aggregation for biofilm formation [106].

Overall, our results suggest which environments one expects virophage to follow a certain mode of coinfection. Overall, virophage that associate with the host are more robust in parameter space and, hence, on average we expect them in a higher diversity of environments. On top of this, space reduces feasibility when virus adsorption rates are high. Because viruses with higher adsorption rates are able to invade systems, this could be detrimental for virophage. Hence, natural mixing of populations through turbulence or some other abiotic means may be necessary for robust virophage coexistence. Further study is necessary to elucidate the impact of space on the full evolutionary dynamics.

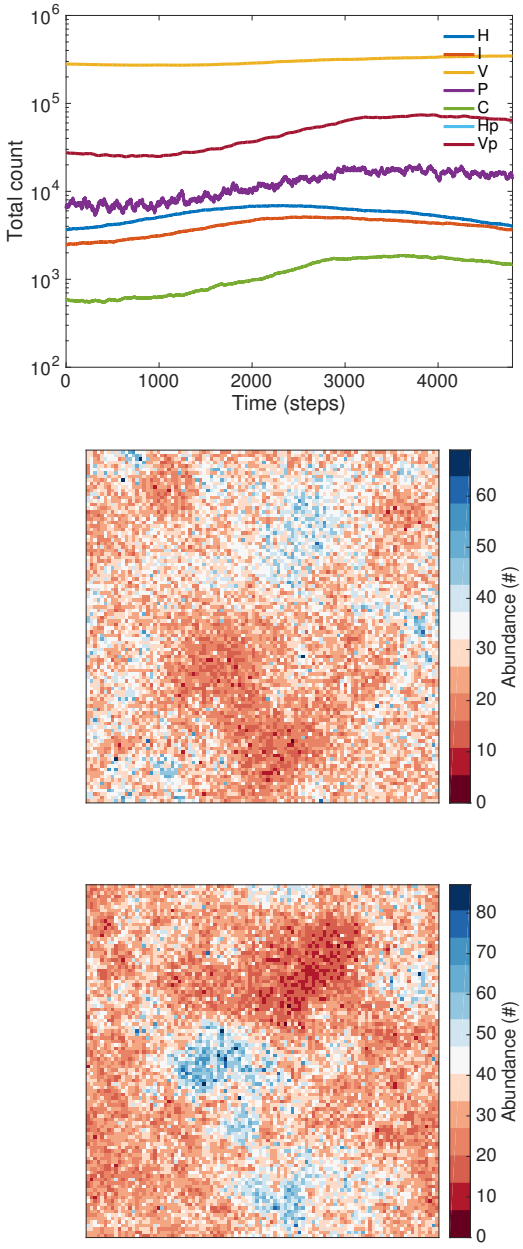


Figure 4.3: Population dynamics associated with selected parameters used for R_{eff} analysis. (Top) Dynamics of total abundance. (Middle) Virus spatial distribution when the population is minimal. (Bottom) Virus spatial distribution when the population is maximal. The colors are chosen to emphasize high and low densities relative to the total range of abundances.

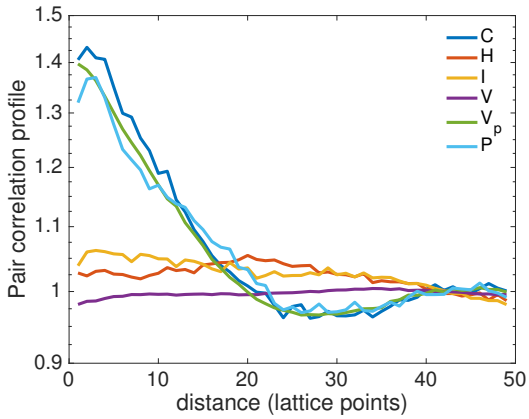


Figure 4.4: Pair cross-correlation profiles of all populations relative to coinfected hosts. A value greater (less) than 1 indicates higher (lower) abundance relative to the expectation from randomly distributing all populations.

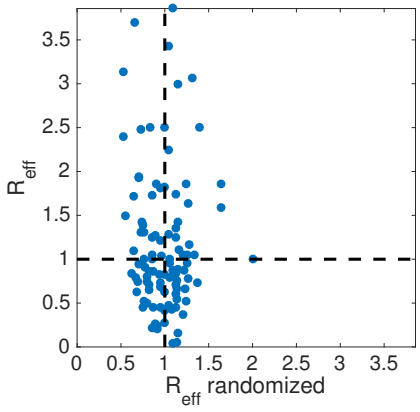


Figure 4.5: Scatter of R_{eff} values for each focal lattice point when considering the observed spatial distribution of populations and when randomizing the spatial distributions. The fitness, R_{eff}, is obtained computationally by tracking the number of secondary lysis events of hosts coinfected as a result of either viral or virophage progeny (explained in more detail in the main text). The demarcating of space by $R_0 = 1$ allows interpretation of the effect of neighborhood on virophage fitness. For example, the top left quadrant indicates points where the local neighborhood increases fitness despite a composition of populations that reduce fitness in the focal lattice point.

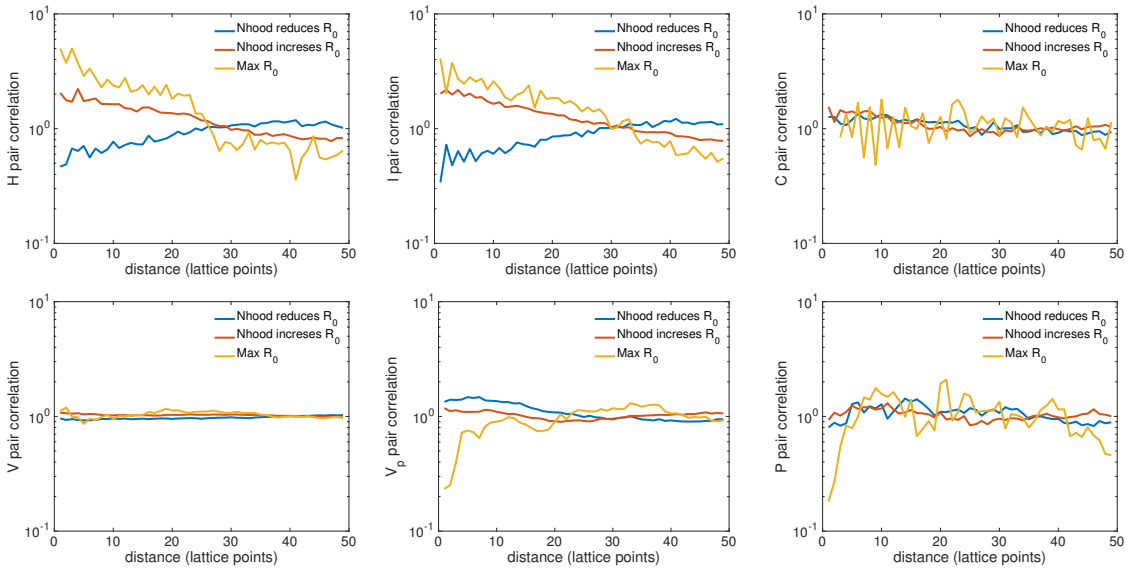


Figure 4.6: Pair cross-correlation profiles of all populations with subsets of the coinfected population. (Top Left) Hosts, (Top Middle) Infected hosts, (Top Right) Coinfected hosts, (Bottom Left) Viruses, (Bottom Middle) Viruses with associated virophage, (Bottom Right) Virophage. A value greater (less) than 1 indicates higher (lower) abundance relative to the expectation from randomly distributing all populations.

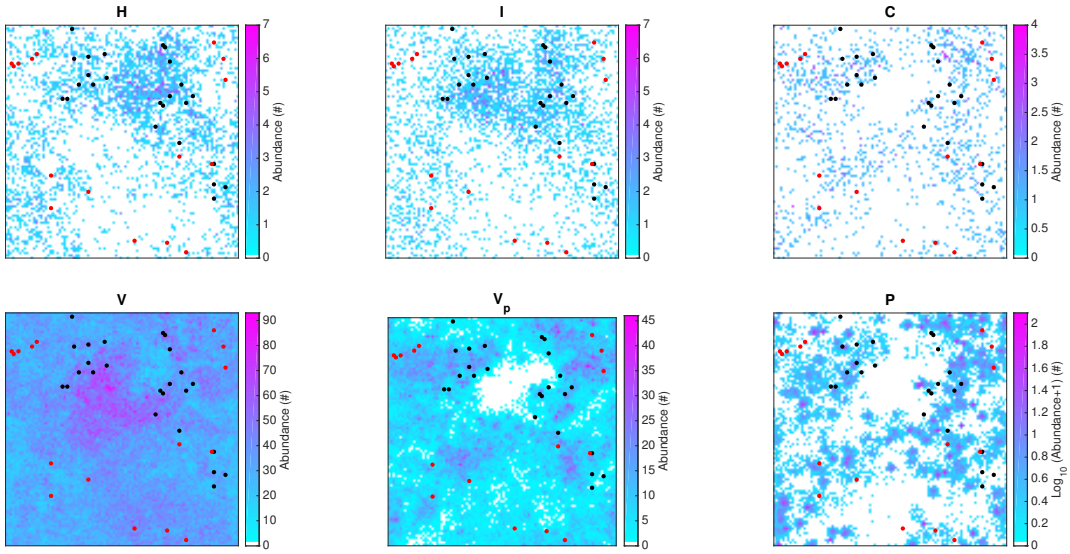


Figure 4.7: Initial state of dynamics with an overlay lattice points that feature strong reliance of fitness on the local neighborhood. Black (red) dots refer to populations where the neighborhood increases (decreases) fitness. (Top left) Hosts (Top middle) Infected hosts (Top right) Coinfected hosts (Bottom left) Viruses (Bottom middle) Viruses with virophage associated (Bottom right) Virophage.

CHAPTER 5
POSSIBLE SHORTENED LATENCY TIMES IN CROV DUE TO MULTIPLICITY INFECTIONS

Broader Context: This chapter addresses what the effect of multiple infections have on infection dynamics. We present data from an experiment where we tracked the infection dynamics of a non-model system: CroV and Cafeteria roenbergensis. We present initial evidence in support of a hypothesis that CroV exhibits reduced burst size with increasing multiplicities of infection. Additionally, we hypothesize a novel viral growth model that features latency times decreasing with MOI. This hypothesis is based on an independence assumption amongst multiple viral factories. We discuss ways to further examine this hypothesis, its applicability to other viruses that utilize viral factories, and its implications for understanding the ecology of giant viruses.

5.1 Introduction

Virus and host life-history traits (e.g., latency time and growth rates) can be estimated from observing infection dynamics. These life-history traits of viruses and their hosts presumably reflect a long-standing coevolutionary dynamic. For example, it is assumed that there exists an optimal viral latency period [107]. This period maximizes viral fitness by trading off the linear production of viruses during infection to the exponential growth in viral populations resulting from new infections in the environment. However, this optimal lysis time varies with standing population of hosts which is, in turn, partially controlled by viral life-history parameters [108]. With this motivation in mind, we seek to empirically characterize life-history parameters across systems – to understand how natural selection guides and is guided by host-viral interactions.

Nucleocytoplasmic Large DNA viruses (NCLDV) are unique in the viral world due to their massive size both physically and genomically [109]. Additionally, many NCLDV reproduce in intracellular compartments of viral origin–virus factories. The virus factories can vary in morphology, composition, and location in the host [26]. Presumably the diversity of viral strategies arise from constraints on the resources provided by host and virus together but also presumably on their effect on the infection dynamics. In order to compare viral production strategies, we must first quantify the infection dynamics of individual systems. A particularly unique viral strategy is the existence of a well-isolated, cytoplasmic viral factory (VF) characteristic to *Cafeteria roenbergensis virus* (CroV) and Mimivirus of *Megaviridae* and Vaccinia virus of *Poxviridae*. These VFs emerge from the "viral core" of an infecting virus. At least for mimivirus and CroV, the viral genome and the subsequent products remain near the VF instead of the common virus tactic of interacting with the host genome [27]. The VFs grow in size as infection proceeds and viruses are produced at the periphery. This virus production tactic is unique in virology from the perspective of multiple viral infections. The separation of viral factories reduces the evolutionary pressures involved with shared viral resources during infection such as the depressor effect [19] and the cheating arising from a tragedy of the commons [20]. First, the number virus factories have been shown to correlate with MOI [27]. Second, entire virus factories can fuse if they are close enough within the cell. Hence, the infection dynamics across regimes of multiplicity of infection (MOI) may vary in unexpected ways.

The infection dynamics of giant viruses, such as NCLDV, across different levels of MOI are largely unknown. To our knowledge, a lone study has investigated the MOI dependence of the infection dynamics between the giant *Aureococcus anophagefferens-Brown Tide virus* and its algal host [110]. That study found a reduction in viral burst sizes at high MOI. This reduction was hypothesized to result from a "depressor effect" due to competition for shared resources [19]. While this virus replicates cytoplasmically in a "viroplasm" there is a lack of evidence suggesting that replication functions similarly to that which oc-

curs in an isolated virus factory [111]. Hence, it is unexplored what effect, if any, MOI has on the infection dynamics of viruses that utilize well-isolated cytoplasmic VFs.

Using CroV, we provide a first in-depth look at the infection dynamics across MOI regimes for giant viruses that utilize organelle-like VFs. We first present the detailed methods outlining how we measured host and viral abundances at hourly periods across 24 hours post infection (hpi). We considered two MOI regimes: a low MOI regime where inoculant virus and host abundances are roughly the same and a high MOI regime with 10 times more viruses. Across the 24 hour period, we see two periods of sustained lysis. From this data we introduce and consider a novel model of viral production appropriate for viruses that utilize well-isolated virus factories. The main prediction of this model is that latency times depend dichotomously on whether a single or multiple viruses infect the host cell. Finally, we estimate relevant life-history parameters including host doubling time, viral burst size, viral DNA production rate, and latency times. We elucidate the effect of MOI on these parameters by comparing between the high and low MOI cases.

5.2 Methods

Suspension cultures of the heterotrophic nanoflagellate *Cafeteria roenbergensis*, strain RCC970-E3 (clonal derivative of strain RCC4623 of the Roscoff Culture Collection), were grown in f/2 artificial seawater medium supplemented with 0.05% (w/v) BactoTM yeast extract (Becton, Dickinson and Company, Germany) as described previously [112]. Exponentially growing cells were diluted with medium to a density of 7.0E5 cells per mL and nine 30 mL aliquots (containing 2.1E7 cells each) were dispensed in 125 mL polycarbonate Erlenmeyer flasks. Three aliquots each served as biological triplicates for the uninfected control cultures, CroV-infected MOI=1 cultures, and CroV-infected MOI=10 cultures. Cultures were infected with *Cafeteria roenbergensis virus* (CroV) strain BV-PW1 [113] at an infectious titer of 1E8 CCID50/mL (the cell culture infectious dose at which 50% of the cultures lyse). Virus titers were determined by end-point dilution assays as described previously

[112].

At t=0 hours post infection (hpi), 210 μL of CroV suspension were added to each of the three MOI=1 cultures, and 2.1 mL of CroV suspension were added to each of the three MOI=10 cultures. Uninfected cultures received 2.1 mL medium. The cultures were then incubated for 15 min at 22°C with gentle agitation. In order to remove any free CroV particles, the cultures were transferred to 50 ml polycarbonate tubes and centrifuged for 10 min at 4,500 rcf, 20°C in an Eppendorf 5804R centrifuge. The supernatants were decanted, the cell pellets were resuspended in 30 mL f/2 medium, and the centrifugation procedure was repeated two more times. After the final wash step, the cells were resuspended in 30 mL f/2 medium with 0.05% (w/v) yeast extract, transferred to 125 mL polycarbonate Erlenmeyer flasks, and incubated at 22°C with 50 rpm shaking.

Aliquots for quantitative polymerase chain reaction (qPCR), flow cytometry (FCM), and microscopy analyses were taken from each culture every hour, starting immediately after infection (0 hpi). To cover the entire infection cycle of CroV, the entire experiment was conducted twice, once to sample from 0 to 9 hpi as well as the 12 and 24 hpi time points, and once to sample from 12 to 24 hpi. Cell concentrations were measured by staining a 10 μL aliquot of the suspension culture with 1 μL of Lugols Acid Iodine solution and counting the cells on a hemocytometer (Neubauer Improved Counting Chamber, VWR Germany). Samples (200 μL) for DNA extraction and subsequent qPCR analysis were processed as described previously [112]. A 128 bp long fragment of the crov283 gene for the VV D11-like transcription factor (GenBank Accession No: ADO67316.1) was amplified by primers CroV-qPCR-9 and CroV-qPCR-10 and used as an approximation for CroV genome copies. PCR conditions have been described previously [112]. For flow cytometry analysis, two 490 μL aliquots (one for FCM analysis, one for backup) were taken per time point and sample and mixed with 10 μl 25% glutaraldehyde in 2 mL cyrovials. After a 20 min incubation at 4°C, the fixed samples were frozen in N_2(l) and stored at -80°C until further analysis. Flow cytometry of viral particles was carried out as described previously [114].

5.3 Results

5.3.1 Infection dynamics

The infection dynamics of CroV on *Cafeteria roenbergensis* over 24 hours post infection (hpi) are shown in Figure 5.1. Hosts were estimated by microscopy counts after Lugol staining and thus include both infected and uninfected hosts. Viral abundance estimates by FCM measure external virus particles. Viral abundance estimates by qPCR measure viral gene copies both inside and outside the cell. There is strong quantitive agreement between replicates (excluding early time viral FCM for MOI=10).

5.3.2 *Cafeteria roenbergensis* growth curve

Host abundances initially decrease between 0 hpi and 2 hpi. This occurs because some hosts will be lost due to the repeated centrifuging and removal of supernatant between 0 and 1 hpi. The decrease between 1 hpi and 2 hpi is unexplained; however, the effect appears to be repeatable as it was observed in both control and non-control replicates. One possible explanation is stress-induced apoptosis through programmed cell death [115] coupled with a lack of growth due to a lag-phase in the growth curve [116]. Both phenomenon have been observed for protist systems, but not explicitly for *Cafeteria roenbergensis*. Hence, we merely note the observation with no claims for the mechanism.

According to standard growth curves, hosts in the control initially grow exponentially following a lag phase. We estimated a host doubling time of 3.27 ± 0.35hr (N=3) using exponential fits to the 3-6 hpi data (Figure 5.2). Our estimated doubling time is longer than previous estimates [117]. For the control, growth slows and eventually saturates by the end of our measurements at 24 hours. For MOI=1, the host abundance increases until 16 hpi when then viral FCM increases significantly. Prior to 16 hpi, the growth rate of the population is less than the control suggesting some combination of restricted growth for infected cells and lysis events. We discuss hypotheses regarding this in the next section.

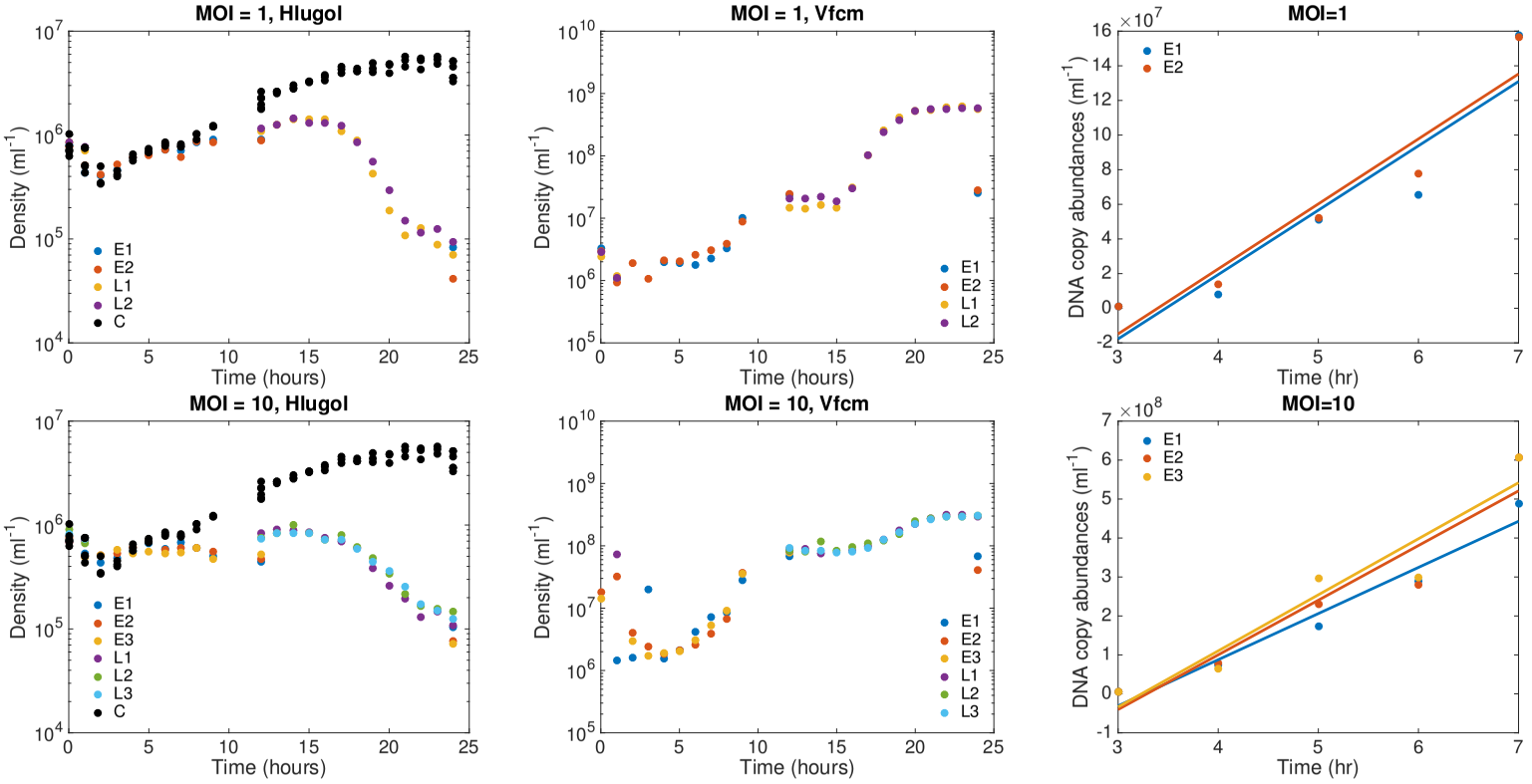


Figure 5.1: Infection dynamics of hosts and viruses. For initial inoculant MOI=1 (Top row) *Cafeteria roenbergensis* abundances from microscopy counts (left), external viral particle abundances from flow cytometry measurements (middle), and viral genome abundances from qPCR (right). For initial inoculant MOI=10 (Bottom row) *Cafeteria roenbergensis* abundances from microscopy counts (left), external viral particle abundances from flow cytometry measurements (middle), and viral genome abundances from qPCR (right). Runs were labeled by whether more of the earlier (E) or later (L) time measurements were included and by the replicate number. All replicates of the the control (C) host abundances without the addition of viruses are plotted in black.

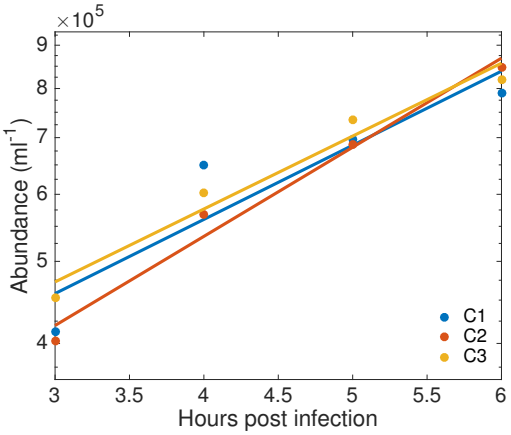


Figure 5.2: Host abundance in control. Data of each replicate is shown as colored dots. Fits of exponential curves used to estimate growth rates are shown as lines.

For MOI=10, the largest decreases in host abundance again begins at 16 hpi. The host abundances decrease between 5-12 hpi suggesting lysis events and cell death occur more frequently than host reproduction.

5.3.3 Hypothesized MOI dependent latency time

The viral FCM data for MOI=1 displays a two distinct periods of increase: the primary lysis event following 6 hpi and the secondary lysis event following 16 hpi. This data is shown next to classic viral growth curves obtained from bacteriophage on *E. coli* in Figure 5.3 [118]. The classic viral growth curves also feature two distinct periods of lysis: the primary lysis event following 28 minutes post infection and the secondary lysis event following 70 minutes post infection. Ellis and Delbruck provided a classical interpretation of the two periods of lysis. They claimed the infections arising from the inoculant viruses leads to the primary lysis event whereas infections arising from viruses produced during the primary lysis event lead to the secondary lysis event. Here, we propose an alternative model of the viral growth curve that may be applicable to CroV and other viruses that utilize cytoplasmic viral factories.

During infection cytoplasmic viral factories produce new CroV particles. Each viral factories grows from the remnant of the core of an individual infecting virus. Thus each

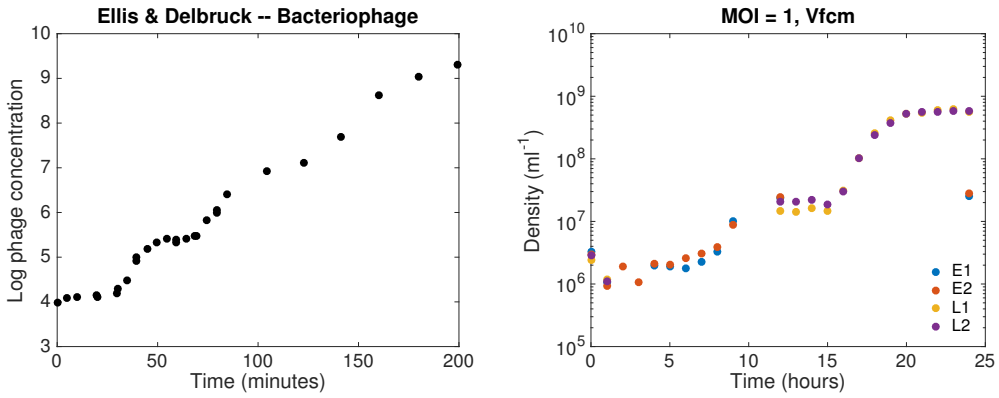


Figure 5.3: Viral growth curves with interpretation of states. (Left) Bacteriophage growth curve on *E. coli* obtained by plot digitizing the data from [118]. All infections resulting from the inoculant viruses cause the primary lysis event between 28 and 60 minutes. The viruses produced from the primary lysis event infect previously uninfected hosts ultimately leading to the secondary lysis event following 70 minutes. (Right) CroV growth curve on *Cafeteria roenbergensis* at MOI=1 obtained from viral FCM. Under our hypothesized growth model, only multiple infections resulting from the inoculant viruses cause the primary lysis event between 6 hpi and 10-12 hpi. A secondary lysis event follows 16 hpi. A majority of the viruses produced during the secondary lysis event result from single MOI infections from the inoculant viruses. Multiple infections resulting from the infections following the primary lysis event also contribute to the secondary lysis event.

infected host is expected to have an increased number of viral factories with increasing MOI (see [27] for the quantitative correlation). It is believed that viral factories are well-contained such that their component parts do not diffuse throughout the cell. Transcription of viral genomes occurs independently of host machinery. Additionally, new virus particles are produced at the periphery of the viral factories. We hypothesize that more viral factories due to high MOI will lead to earlier lysis times. The underlying assumption is that virus factories proceed to lysis following exhaustion of a host resource and that the rate of utilization of this host resource grows linearly with the number of viral factories.

An idealized case of infection dynamics between MOI=1 and MOI=2 is shown in Figure 5.4. Each host is infected at the same moment and both infection proceeds through an eclipse time, t_ε. By this point only viral transcription has initiated and no full virus particles have been produced. Following the eclipse time viral particles are produced at the periphery of the viral factories. Under our assumption, the limited host resources cause

Table 5.1: Models of MOI effect on single infection dynamics. The invariant model assumes no effect of MOI on lysis time, τ, and burst size, β. The other models are presented in comparison to the invariant models. Only lysis times vary across models.

Model	latency time	burst size	Reference
Invariant	τ	β	-
Accelerated	$< \tau$	$\approx \beta$	this work
Delayed	$> \tau$	$\approx \beta$	[119]

the viral burst size in both cases to be the same. However, the MOI=2 case produces viral particles at twice the rate due to twice the number of viral factories. Thus following the eclipse time, which is the same in both cases, the MOI=2 infection lyses in half the time and with the same burst size as the MOI=1 infection. A comparison of our predictions to other viral growth models is presented in Table 5.1.

Here we develop a dynamical model of infection dynamics in order to quantitatively compare our hypothesized accelerated lysis time model to an invariant lysis time model. Hosts are initially uninfected or susceptible (S) and become exposed (E) following adsorption of viruses. Exposed hosts, in a dynamical sense, are unaffected by viral infection. That is, exposed hosts still adsorb viruses leading to multiple infections. Similarly, exposed hosts implicitly consume resources and contribute to the density dependent growth of susceptible hosts. Exposed hosts eventually transition to an infected (I) state following an average eclipse time. Infected hosts no longer adsorb viruses and, concurrently, no longer consume resources or reproduce. Infected hosts eventually lyse after an average lysis time that may be MOI dependent depending on the model. Lysis removes the infected host and adds a burst of viruses. The model can be described mathematically as:

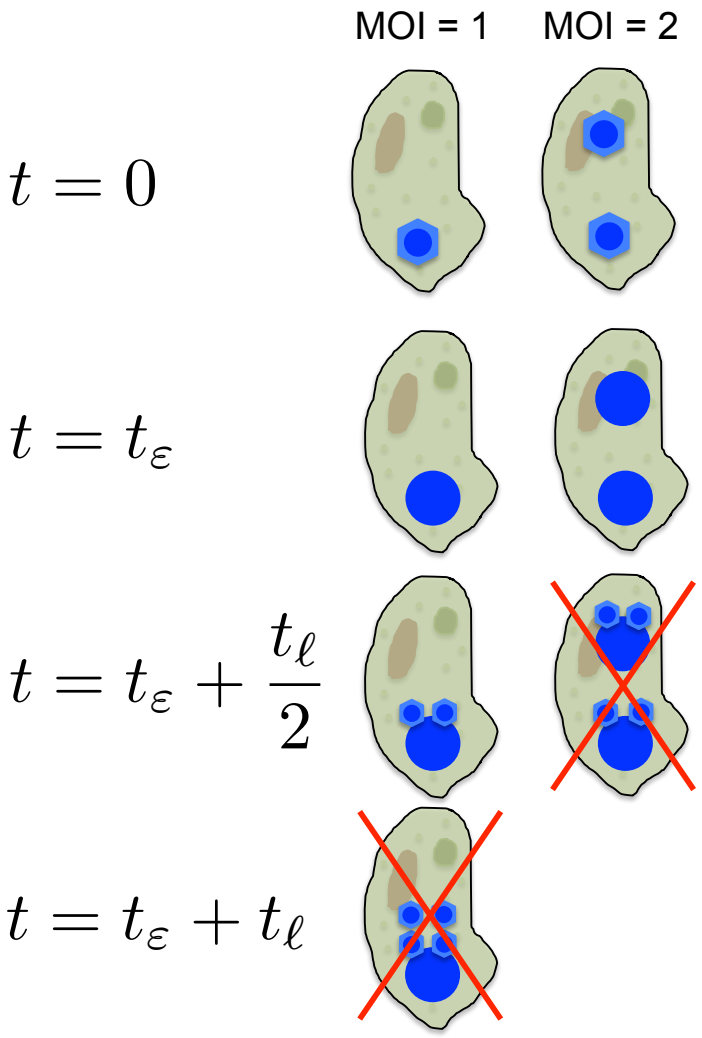


Figure 5.4: Hypothesized MOI-dependent infection dynamics due to viral growth using viral factories. Two identical hosts are infected by CroV with different MOI at $t = 0$. Viral transcription initiates and the viral factories grow in size throughout the eclipse period t_ε. Viruses are produced at the periphery of the viral factories. In this example, lysis occurs after the host resources due to the production of the burst size, $\beta = 4$, of new virus particles. Following the eclipse period, it takes half the time to produce the burst size for the MOI=2 case as compared to the MOI=1 case.

$$\dot{S} = r(S + \sum_{ij} E_i^j)\left(1 - \frac{H + \sum_{ij} E_i^j}{K}\right) - \phi SV \tag{5.1}$$

$$\dot{E}_1^1 = \phi V(S - E_1^1) - \eta_0 E_1^1 \tag{5.2}$$

$$\dot{E}_i^j = \phi(E_{i-1}^j - E_i^j)V + \eta_0(E_i^{j-1} - E_i^j) \tag{5.3}$$

$$\dot{I}_i^1 = \eta_0 E_i^n - \eta_i I_i^1 \tag{5.4}$$

$$\dot{I}_i^j = \eta_i(I_i^{j-1} - I_i^j) \tag{5.5}$$

$$\dot{V} = \beta \sum_i \eta_i I_i^{m_i} - \phi V(S + \sum_{ij} E_i^j) \tag{5.6}$$

where subscripts refer to MOI and superscripts refer to different stages over time within the exposed class and the infected class. The use of multiple stages over time within one state is common in epidemiological modeling and is often referred to as "boxcars" [120]. The dynamical result of staged transitions is the switching time between states is gamma distributed. The rates of transition between stages η is the same across MOI during the eclipse period and can vary across MOI during the latency period. The number of stages is n for the exposed state and m_i for the infected state. The number of stages sets the number of summations over j in the model. With fixed transition rates, increasing the number of stages reduces the variance of the respective gamma distribution. In the model, r is the growth rate of hosts, K is carrying capacity or maximum abundance of hosts possible, and ϕ is the adsorption rate of viruses. We assume that viral decay is negligible over the 24 hour period. A graphical interpretation of the mathematical model is shown in Figure 5.5.

Figure 5.6 shows example MOI=1 dynamics from the ODE model across a range of adsorption rates. The left columns shows viral growth curves when the average lysis time, t_ℓ, following the eclipse period is set to 5 hours. The right columns shows viral growth curves when the average lysis time, t_ℓ, after the eclipse time is maximally 14 hours and decreases harmonically with MOI (e.g., $t_\ell = 7$ hours for MOI=2). The MOI is capped to MOI=4. All gamma distributions are set to have a variance of 1 hour. The eclipse

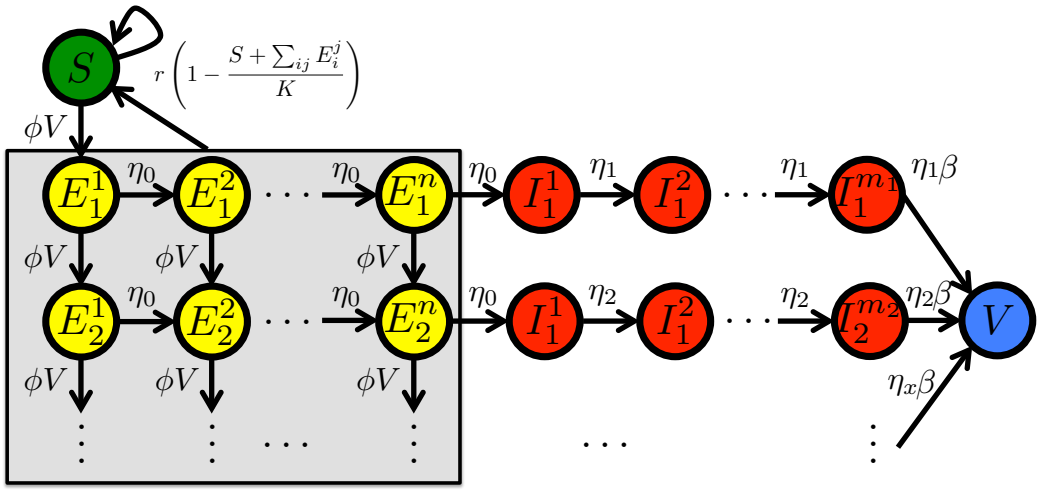


Figure 5.5: Transitions between states of host from infection to lysis. Susceptible hosts transition to an exposed state E_1^1 following viral adsorption at rate ϕV. Subsequent viral adsorption to exposed hosts occur at rate ϕV and increases the value of the subscript by 1. Exposed hosts transition to later stages of exposure at a rate η_0 leading to a gamma distribution of times to transition from exposed to infected I. All susceptible and exposed states of host increase the susceptible population through reproduction at the density dependent rate $r\left(1-\frac{S+\sum_{ij}E_i^j}{K}\right)$. Infected hosts transition to later stages at a rate η_x where x may depend on the multiplicity of infection (MOI). Infected hosts produce β viruses as a result of lysis. Lysis follows the last stage of infection, m_x, which may depend on MOI=x.

time for both models is gamma distributed with an average period of 4 hours. We set the initial condition to be a poisson distributed across allowable MOI according to the inoculant MOI. For both models, the secondary lysis event occurs earlier in time as viral adsorption increases. The same results for the MOI=10 case are shown in Figures 5.7.

We find a consistent signal: the growth curves between models are qualitatively similar and difficult to differentiate quantitatively. In particular, the quality of fit depends on the unknown adsorption rate of the virus. Both models feature dynamics consistent with the data.

5.3.4 CroV Burst size

A standard method for estimating the burst size of virus is performing plaque assays at varying dilutions. Unfortunately, currently there are no methods for producing lawns of *Cafeteria roenbergensis*. Instead we estimate the viral burst size from the changes in the host and virus populations during a lysis event. We only consider the lysis event after 15 hpi because the magnitude of decreases in the host population is large. This reduces impact of noise arising from host growth during the lysis events. We estimate the burst size, $\beta = \frac{V(t=24)-V(t=15)}{H(t=15)-H(t=24)}$, using the FCM data. The burst size is larger for MOI=1 with $\beta_{\text{MOI=1}} = 440 \pm 23.8$ (N=2) than for MOI=10 with $\beta_{\text{MOI=10}} = 311 \pm 5.3$ (N=3), where the $\pm$ refer to standard error here and throughout this chapter–error propagation is performed, when possible. According to both model hypotheses, we expect that the secondary lysis event features more multiple infections in the MOI=10 case versus the MOI=1 case. Thus we conclude that the burst size decreases with increasing MOI.

To further address the role of MOI dependence on the burst size we estimate the burst size of the primary lysis event. However, we can not directly measure the change in the host population due to lysis because of the concurrent host growth. However, if assume well-mixed dynamics and total adsorption of viruses during the inoculation, we can leverage expectations from the Poisson distribution to estimate the burst size. For MOI=10, the

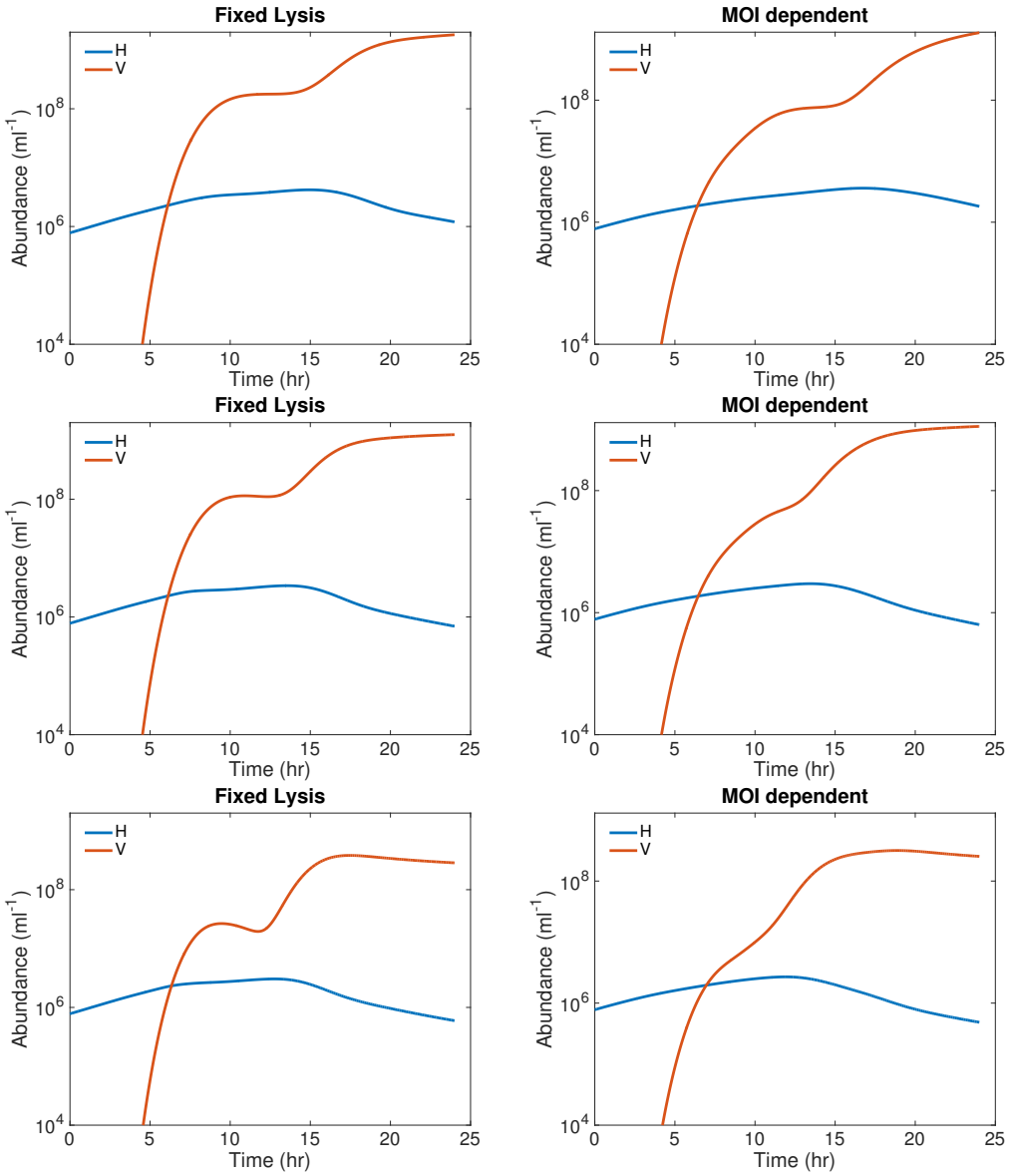


Figure 5.6: Simulated viral growth curves for MOI=1. (Left Column) Model where latency times do not depend on MOI (Right column) Our proposed model where latency times decrease harmonically with MOI. The adsorption rates increase going down with (top row) $\phi = 9.17E-9$ ml hr^{-1}, (middle row) $\phi = 9.17E-8$ ml hr^{-1} and (bottom row) $\phi = 9.17E-7$ ml hr^{-1}. Other parameters are host growth rate, $r = 0.21$ hr^{-1}, carrying capacity, $K = 4E6$ ml $^{-1}$ and burst size, $\beta = 400$. The initial host population is $H_0 = 7.8E5$ ml^{-1}. The initial host population is infected according to a poisson distribution parametrized by the MOI.

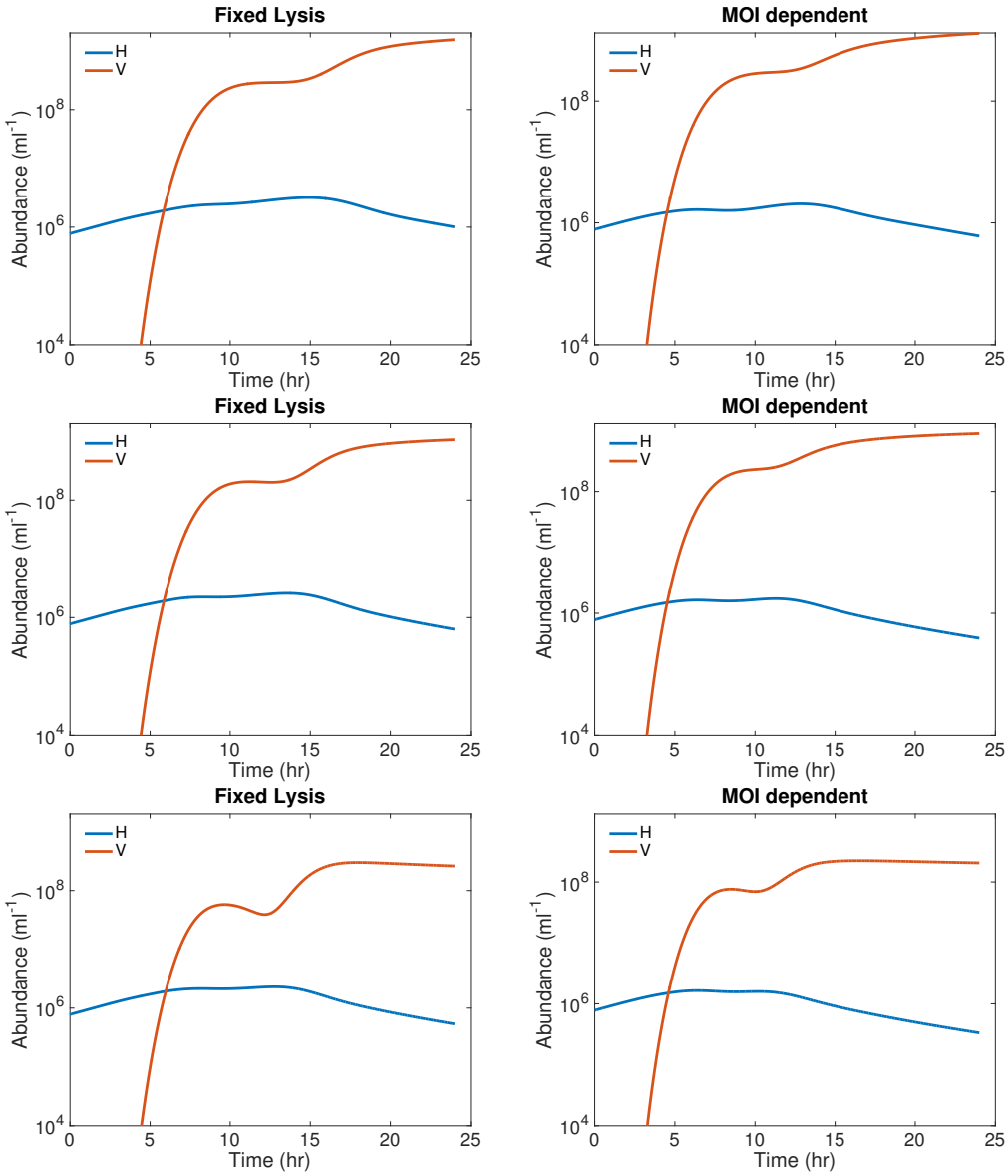


Figure 5.7: Simulated viral growth curves for MOI=10. (Left Column) Model where latency times do not depend on MOI (Right column) Our proposed model where latency times decrease harmonically with MOI. The adsorption rates increase going down with (top row) $\phi = 9.17E-9$ ml hr^{-1}, (middle row) $\phi = 9.17E-8$ ml hr^{-1} and (bottom row) $\phi = 9.17E-7$ ml hr^{-1}. Other parameters are host growth rate, $r = 0.21$ hr^{-1}, carrying capacity, $K = 4E6$ ml $^{-1}$ and burst size, $\beta = 400$. The initial host population is $H_0 = 7.8E5$ ml^{-1}. The initial host population is infected according to a poisson distribution parametrized by the MOI.

Poisson distribution expects 99.95% of the host inoculant to be singly or multiply infected. This allows estimating the burst size: $\beta = \frac{V(t=12)-min(V)}{H(t=2)}$ where we take the minimum of V given noise the viral FCM data and we considered the host population following the decline at the beginning. This yields $\beta = 158.8 \pm 3.85$ for MOI=10. Note, this value is considerably smaller than previous estimates from the secondary lysis events. Two reasons may account for this: MOI dependence of the burst size and viral adsorption onto host cells during the primary lysis event. For MOI=1, the Poisson distribution expects 63% of inoculant host cells to be infected by a virus. This scales our burst size $\beta = \frac{V(t=12)-min(V)}{(1-e^{-1})H(t=2)}$ yielding $\beta = 88.7 \pm 0.8$. This contradicts our expectation because the case with lower MOI has a lower burst size. However, if we assume that only multiply infected cells lyse during the primary lysis event we obtain a different scaling factor $\beta = \frac{V(t=12)-min(V)}{(1-2e^{-1})H(t=2)}$ yielding $\beta = 177.46 \pm 1.63$. That is the early time burst sizes are more comparable when using only multiply infected host cells to estimate the burst size. The burst size is still modestly larger in the MOI=1 case suggesting that there may be MOI dependence on the burst size.

5.3.5 Viral production rate from viral factory

The viral qPCR data shows that intracellular viral DNA replication initiates by 3 hpi. Interestingly the final yield of viral genomes is approximately the same with V_{PCR}(t=24) = 5.59E8$\pm$2.15E6 for MOI=1 and V_{PCR}(t=24) = 5.62E8$\pm$8.17E6 for MOI=10. This suggests the total host resources limit total viral production. In particular, the MOI=10 case saturates to the final yield much more quickly (7 hpi versus 12 hpi).

One prediction of our hypothesized model is the growth rate of viral genomes should be linearly proportional to the number of viral factories. We estimate the growth rate of viral DNA copies per virus factory by first estimating the intracellular viral production rate of the population from the slope linear fit to the PCR data between 3 and 7 hpi. We divide this by the viral abundance at 2 hpi. The estimated viral production rate is 75 DNA copies/hr/viral factory for both MOI=1 (76.4 ± 30.0, N=2) and MOI=10 (73.5 ± 35.8, N=3). The larger

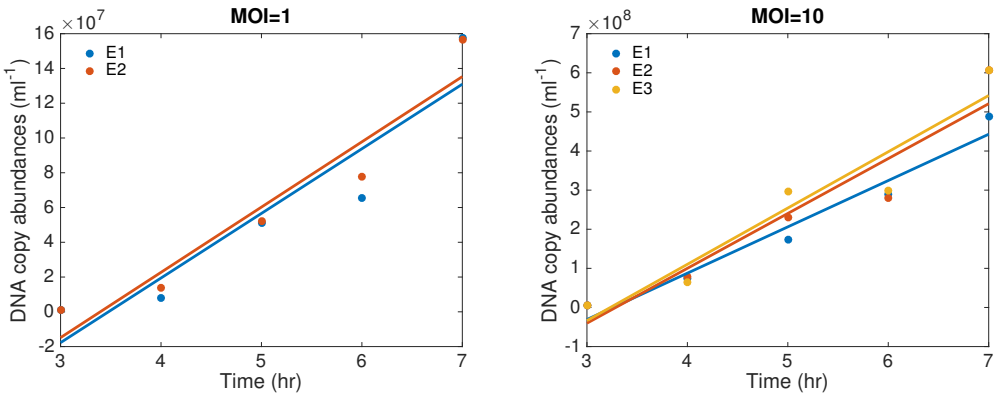


Figure 5.8: Early-time growth of intracellular viral DNA copies from the qPCR data. (Left) MOI=1 (Right) MOI=10. Linear fits to the data are shown as lines. The growth rates associated with these lines were used to estimate the average growth rate of viral genomes per viral factory.

standard error for the MOI=10 case may be the result of virus factories coalescing in high MOI infected hosts cells. Note, the rate of gene copies decreases following this period likely due to an exhaustion of host resources and the subsequent lysis of infected cells.

5.4 Discussion

We characterized the infection dynamics across high and low MOI of CroV on *Cafeteria roenbergensis*. We counted host cells by microscopy and viral abundances by both FCM and qPCR. We estimated a *Cafeteria roenbergensis* doubling time of 5.5 hours. Based on the known biology of viral factories, we proposed a novel model of viral growth leading to shortened latency times due to multiple infections. The model is based on two assumptions: 1) the intracellular viral growth rate increases linearly with the number of viral factories and 2) lysis occurs once a fixed burst size is reached.

Using simulations we attempted to discern our model to one where latency times were fixed. However uncertainty in the adsorption rate led to no discernible differences between the two models. At this time, we cannot yet eliminate either the classic nor this new model as the underlying mechanism. Nonetheless, there is some preliminary support for reduced latency times when estimating the effect of MOI on burst sizes. We estimated the burst size

of the second period of lysis which led to a burst size of 440 viruses for the MOI=1 case. The MOI=10 was lower suggesting either changes to the infected host growth rates or to MOI dependency of the burst sizes. We justified our model by showing estimates of the primary lysis burst sizes match when considering only multiple infections. Additionally, our hypothesis predicts that all virus factories are initially independent. This assumption led us to identifying a linear rate of 37.5 DNA copies/hr/virus factory following a viral eclipse period of 3 hpi. We also estimated from the control data a host doubling time of 3.27 hours.

Despite the early growth rate of DNA copies this does not necessitate that the intracellular viral production follows the same rate. Viruses form at the periphery of the virus factory [27]. Thus, we expect the intracellular viral growth rates to scale with the surface area of the virus factories while the DNA production scales with the volume. Our model with still predicts a viral growth linear with the number of viral factories. However, this discrepancy in rates may account for MOI dependence in burst size. Additionally, if virus factories must interact with the host nucleus via transcription factors [112]. The diffusion of these proteins and others related to production may lead to different viral factories initiating later stages viral production earlier. This could partially explain the observed MOI dependence of burst rates. Similar MOI dependence has been shown in a virus system that utilizes some form of viroplasm for reproduction [110].

The presentation of our hypothesized model focused on an ideal scenario ignoring interactions between viral factories. Rather, it has been argued that virus factories are able to fuse if they are proximate enough [27]. This would lead to non-linearities in our proposed model. The infection dynamics could be considered instead from the perspective of MOI to effective MOI based on the number of independent viral factories in a cell. This would complicate the dynamics such that the adsorption rate would also be dependent on what the actual MOI is. For example, the adsorption rate for the initial infection would be larger than the adsorption rate for the secondary infection because two viral factories

might require more than 2 viruses to adsorb. Hence, we do not claim that our model is fully representative of the actual infection dynamics. Rather, we suggest this model provides an appropriate alternative hypothesis upon which to build to consider further complexity. Further study is warranted to determine the validity of this model. Additionally, comparison across systems should be done to determine whether this model is general and applied to other viral factory utilizing viruses such as *Mimivirus*.

CHAPTER 6
CONCLUSIONS

In this thesis we studied the population dynamics of multiple infection in viral systems. We focused on addressing both the causes in consequences of multiple infection in order to provide a fuller context for how multiple infections may feedback to affect the ecology and evolution of viruses and their hosts. A particular focus of this thesis was on the role of multiple infections and coinfection in under-explored viral systems such as virophage and nucleocytoplasmic large DNA viruses. The global impact of these viruses has yet to be fully understood [121]. This thesis represents a step forward in understand the ecology of these viral entities and it suggests that multiple infections play a major role.

Virophage dynamics constitute a central motivation for this thesis. These viruses of viruses are unique in the viral world for existing in tri-partite relationships with hosts and viruses. A reliance on coinfection has been characterized in other viral systems. Segmented viruses each contain a portion of the entire viral genome needed for replication [68], satellite viruses rely on a helper virus so heavily that it does not encode for its own capsid proteins [122], and defective interfering particles represent cheats that derive from full viruses following serial infections at high multiplicity of infection [123]. Nevertheless, the relationship virophage share with their hosts and viruses is unique. They exist as independent viruses that parasitize a unique niche. For this reason, they seemingly face little competition from other viruses. Additionally, since they function similarly to a group level immune system from the giant viruses there is strong pressure for the hosts to maintain the virophage population. Indeed, recent work has shown lysogenic like virophage which can reproduce vertically in the host population [112]. Nevertheless, virophage still rely on the viruses upon which they hyper-parasitize. Understanding the full complexity of natural selection in this tri-partite system should be explored in future studies.

The study of virophage remains in its infancy. Currently, a majority of the virophage have been identified indirectly from metagenomic datasets [39]. The diverse areas in which virophage are found are particularly impressive given that we only have two main virophage signatures in gene databases: the sputnik capsid protein and the mavirus capsid protein. Presumably, additional isolation of virophage will yield new signatures by which to identify these elusive hyperparasites. Isolating novel virophage, much like the virophage existence itself, is part of a larger hyper-problem–isolating giant viruses. This begs the further question, what is the ubiquity of giant viruses that utilize cytoplasmic virus factories? Much of our knowledge regarding giant viruses has been limited due to filtering practices typical to sampling marine environments [124]. The filters commonly used to isolate marine viruses are too small for giant viruses to pass through. Additionally, isolating the virus relies on identifying its host which is a further compounded by developing culturing conditions using the hosts. Hopefully, the notion of pioneering an area of virology will entice more researchers to explore the giant virus and virophage world further.

I am not that bold. However, Matthias Fischer is and our collaboration gave me the opportunity to include an experimental component to this thesis. We addressed how multiplicity of infection affects the infection dynamics of the giant virus CroV. Testament to the vast unknown in the giant virus world, we proposed a novel dynamic in virology: latency times that decrease with multiplicity of infection. The data is fairly coarse, despite a high frequency of samples across time. The bottleneck in the process is measurement. Host counts can only be done reliably by microscopy and even the viral FCM data featured noise in certain regimes. This is far from the high-throughput methods afforded to researchers of model systems. The risks in studying a non-model system may grow as biology becomes more quantitative and larger data-sets are expected for publication. Our work and preliminary finding warrants follow-up investigation.

This thesis identified causes of multiple infection (e.g., spatial clustering) along with empirical effects (e.g., dichotomous lysis times). Future work on the theory side involves

understanding how the dynamics feedback. Do MOI dependent lysis times increase or decrease the frequency of multiple infections? This fits in a broader conversation of optimal lysis times [108]. Earlier MOI lysis are counter to what's expected to be optimal. Is this a mechanistically derived detriment that all giant viruses that utilized well-isolated viral factories must face? Or is there a theoretical benefit for such a strategy? Perhaps this may suggest that these viruses only occur in low-density environments where multiple infection is unlikely. Perhaps, the lower burst sizes resulting from multiple infection provides a means of regulating the viral population to an optimal level. Furthermore, it would be particularly interesting to ask whether viruses can cheat by varying the properties of their virus factory. Would faster reproducing virus factories be analogous to some sort of prisoner's dilemma scenario? Theoretical models could shed insight into these questions.

Overall, this thesis sheds light on the importance and argues for the ubiquity multiple viral infections of microbes. The full scope of the dynamical role of multiple infections play in situ has yet to be elucidated; however, we hope this work will drive the community towards further investigations.

APPENDIX A

SUPPLEMENTARY MATERIALS FOR CHAPTER 2

A.1 Mean-field steady state

The dynamical system is:

$$\frac{dH}{dt} = rH\left(1 - \frac{H + \sum_j I_j}{K}\right) - dH - \phi HV \tag{A.1}$$

$$\frac{dI_1}{dt} = \phi(H - I_1)V - (d + \lambda)I_1 \tag{A.2}$$

$$\frac{dI_k}{dt} = \phi(I_{k-1} - I_k)V - (d + \lambda)I_k \tag{A.3}$$

$$\frac{dV}{dt} = \beta\lambda \sum_j I_j - \phi(H + \sum_j I_j)V - mV \tag{A.4}$$

We set the left-hand-side equal to 0 and solve for the equilibrium value of the variables. They are denoted with a *. From (A.2) we get:

$$I_1^* = \frac{\phi H^* V^*}{\phi V^* + \lambda + d} \tag{A.5}$$

Combining this result with (A.3) gives the geometric sequence for the infected classes:

$$I_k^* = H^*\left(\frac{\phi V^*}{\phi V^* + \lambda + d}\right)^k \tag{A.6}$$

Thus $H^* + \sum_j I_j^*$ is an infinite geometric series that converges to:

$$H^* + \sum_j I_j^* = H^* \frac{1}{1 - \frac{\phi V^*}{\phi V^* + \lambda + d}} = H^* \frac{\phi V^* + \lambda + d}{\lambda + d} = H^*\left(1 + \frac{\phi V^*}{\lambda + d}\right) \tag{A.7}$$

Substituting this into (A.1) yields:

$$0 = rH^*\left(1 - \frac{H^*}{K}\left(1 + \frac{\phi V^*}{\lambda + d}\right)\right) - dH^* - \phi H^* V^* \implies$$
$$H^* = \frac{K}{r}\frac{(\lambda + d)(r - d - \phi V^*)}{\lambda + d + \phi V^*} \tag{A.8}$$

when $H^* \neq 0$. Substituting the results of the geometric series into (A.4) yields:

$$0 = \beta\lambda H^* \frac{\phi V^*}{\lambda + d} - \phi V^* H^*\left(1 + \frac{\phi V^*}{\lambda + d}\right) - mV^* \implies$$
$$V^* = \frac{1}{\phi}\left(\beta\lambda - (\lambda + d) - \frac{m(\lambda + d)}{\phi H^*}\right) \tag{A.9}$$

Substituting (A.9) into (A.8) yields:

$$H^* = \frac{K}{r}\frac{(\lambda + d)(r - d - \left(\beta\lambda - (\lambda + d) - \frac{m(\lambda+d)}{\phi H^*}\right)}{\left(\lambda + d + \left(\beta\lambda - (\lambda + d) - \frac{m(\lambda+d)}{\phi H^*}\right)\right)} \implies$$
$$H^*\left(\beta\lambda - \frac{m(\lambda + d)}{\phi H^*}\right) = \frac{K}{r}(\lambda + d)\left(r - \beta\lambda + \lambda + \frac{m(\lambda + d)}{\phi H^*}\right) \implies$$
$$r\phi\beta\lambda(H^*)^2 - (\lambda + d)(mr + K\phi(r - \lambda(\beta - 1)))H^* - Km(\lambda + d)^2 = 0 \tag{A.10}$$

Since our parameters our non-negative, there are real solutions to this quadratic. Since $Km\phi r\beta(\lambda + d)^2 > 0$, one solution is negative and can be ignored. The simplified, positive H^* steady state is.

$$H^* = \frac{\lambda + d}{2r\phi\beta\lambda}\left(mr + \phi K(r - \lambda(\beta - 1)) + \sqrt{(mr + \phi K(r - \lambda(\beta - 1)))^2 + 4Km\phi r\beta\lambda}\right) \tag{A.11}$$

Subsequently,

$$V^* = \frac{1}{\phi}\left(\beta\lambda - (\lambda + d) - \frac{2mr\beta\lambda}{mr + \phi K(r - \beta\lambda + \lambda) + \sqrt{(mr + \phi K(r - \lambda(\beta - 1)))^2 + 4Km\phi r\beta\lambda}}\right) \tag{A.12}$$

A.2 Viral clustering due to varying D_v

The top two rows of Figure A.1 show snapshots of the spatial dynamics of hosts and viruses respectively for different values of D_v while keeping all other parameters constant. Spatial clustering increases with decreasing D_v (columns moving left to right). The corresponding radial pair correlation profiles with the threshold value of 1.1 are shown in the bottom row of Figure A.1. High D_v dynamics lead to cluster profiles indistinguishable from random dispersal of hosts and viruses (left column). Reducing D_v increases clustering of the viruses, but not hosts (middle and right columns).

As in the main text, we show spatial clustering increases the severity of MIs when we vary D_v. The top rows of Figure A.2 show the observed MOI distributions as compared to the previously derived mean-field expectation (black line) as we vary D_v. We normalize the distributions by setting the density of singly infected hosts to 1. The observed MOI distributions match the mean-field geometric sequence for high D_v. Spatial clustering leads to an increase in the mass of the tail of the MOI distribution, i.e., more hosts are infected by more viruses as compared to mean field.

The MOI distributions have relatively "fat" tails because more viruses are colocated with hosts of increasing MOI. The bottom row of Figure A.2 shows the probability distributions of observing a number of external viruses in lattice points that contain hosts with a specific MOI. For clarity we only show these viral probability distributions (VPDs) up to MOI=6. Randomly distributing hosts and viruses across the domain leads to Poisson distributed VPDs (black lines) parametrized by the observed viral density. The observed VPDs match the Poisson distribution for all MOI in the high D_v. Whereas for all other

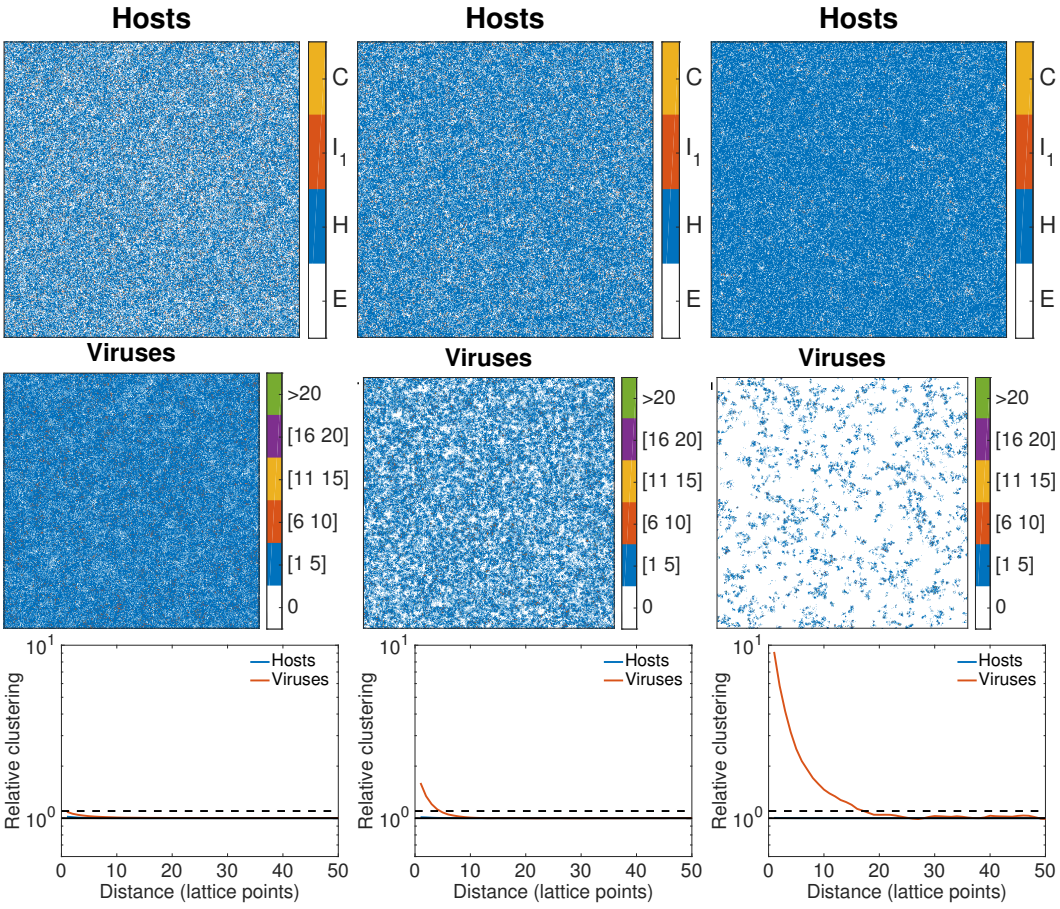


Figure A.1: Spatial clustering of viruses increases with decreasing the viral diffusion constant. The color legend for hosts refers to multiply infected hosts (C), singly infected hosts (I_1), uninfected hosts (H) and empty lattice sites (E). The color legend for viruses refers to the number of viruses located at each lattice point. (Left Column) High viral diffusion, $D_v = 2 * 10^{-4.5}$ (cm^2/hr). (Middle Column) lowered viral diffusion, $D_v = 2 * 10^{-5.5}$ (cm^2/hr) (Right Column) further lowered viral diffusion, $D_v = 2 * 10^{-6.0}$ (cm^2/hr). Rows correspond to (Top) distribution of hosts, (Middle) distribution of viruses and (Bottom) radial pair correlation profile of hosts. The dotted line at $y = 1.1$ approximates the 99% confidence interval of the pair correlation profile when hosts and viruses are randomly dispersed. We use the intersection of this threshold line and the observed pair correlation profiles to define the cluster widths. When clustering occurs the corresponding cluster widths are plotted as black lines outside the top left corner of each of the spatial distribution plots. Other parameter values are shown in the table in the appendix of the main text.

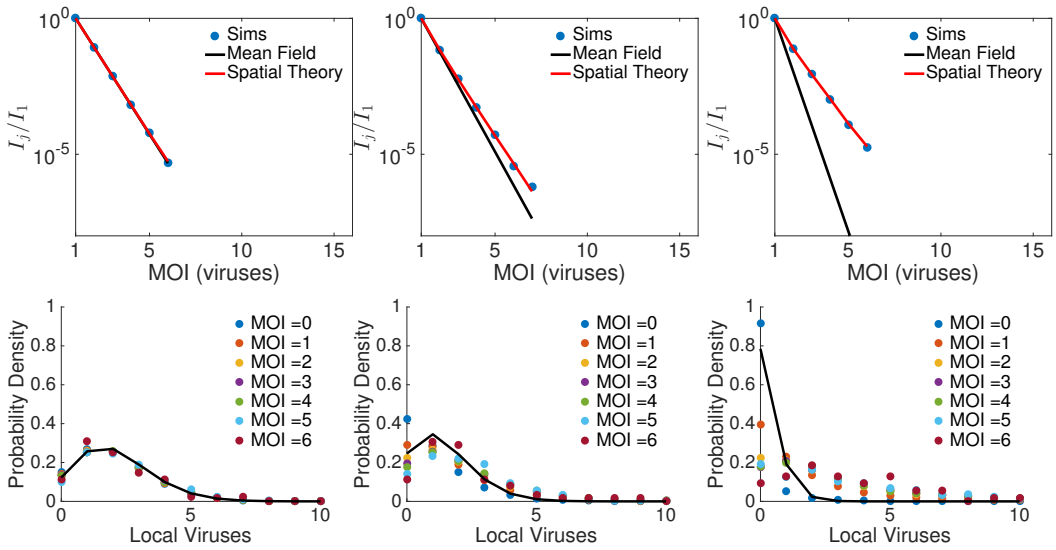


Figure A.2: MOI distributions arising from local densities of viruses and underlying non-Poisson distributions of VPDs. Predictions from spatial theory obtained using MOI distribution predictions in main text. (Top) MOI distributions and (Bottom) distributions of colocated viruses conditioned on MOI host for (Left Column) high viral diffusion, $D_v = 2.04 * 10^{-4.5}$ (cm^2/hr). (Middle Column) lowered viral diffusion, $D_v = 2.04 * 10^{-5.5}$ (cm^2/hr), and (Right Column) further lowered viral diffusion, $D_v = 2.04 * 10^{-6.0}$ (cm^2/hr). Other parameters are the same as in Figure A.1

cases, the observed VPDs deviate from the Poisson distribution by skewing to the right, i.e., there are more viruses colocated with high MOI hosts than expected in the mean field theory. This skewing is more pronounced as clustering increases. In contrast, the VPD for uninfected hosts skews to the left due to viral clustering for low D_v. This skewing results from the emergent, temporary virus-free domains.

A.3 Spatial model processes

The spatial model updates at fixed time steps where processes occur stochastically. The processes and associated parameters are listed in Table A.1. The adsorption rate, ϕ, and viral diffusion constant, D_v, are varied in our analysis. When one parameter is varied the other is fixed to the value presented in the table. All processes are Poisson. Hence, the length of time, τ, for a single time step dictates the conversion of rates to probabilities via an exponential distribution: $P_{\text{event}} = 1 - e^{-r\tau}$ for some rate r.

We assume that adsorption of the viral particles are diffusion limited and we solve for the rates following [126]. A spherical cell's intake of spherical ligands follows:

$$J = \ 2\pi Dac_\infty$$

where J is the maximum rate of absorption of ligands with diffusion constant D and far-off concentration c_∞. The absorbing spherical cell has diameter a. Our models use the adsorption rate, ϕ calculated from J:

$$\phi = \frac{J}{c_\infty} = 2\pi Da,$$

The diffusion constant, D, is estimated from the Stokes-Einstein Relation: $D = \frac{k_b T}{3\pi\eta d}$. This relation is relevant for spherical particles in low reynolds number fluids which is typical at micro-organismal length scales. This leads to:

$$\phi = \frac{2ak_bT}{3\eta d},$$

for a spherical molecule with diameter d in a fluid with viscosity η at temperature T where k_b is the Boltzmann constant..

Diffusion is simulated by individuals probabilistically hopping to an adjacent lattice-point (including diagonals) within a time step. Diagonal hops occur less frequently in order to maintain the spatial isotropy of diffusion. Diffusion constants are based on physical models dictated by the size of the individual and viscosity of environment. We choose τ such that hopping to a lattice point outside of the neighborhood has less than 5% chance per time step, according to a physical diffusion model for movement. The probability for a virus to move adjacently is: $P_{\text{move}} = 1 - p_0^2$ where $p_0 = \sqrt{1 - \frac{4D_v\tau}{\Delta x^2}}$. Host movement occurs regardless of infected state. Attempted moves into occupied lattice points are aborted in

order to maintain the single host occupancy rule. Similarly, uninfected hosts reproduce respecting the same rule.

Host death and viral decay remove individuals from the lattice. Viral infection can occur when viruses share a lattice point with a host, i.e., a host and virus are colocated. Each colocated virus has the same probability of infecting the host as determined by the adsorption rate, ϕ. Hence, multiple viruses can infect the same host in one time step following binomial distributions. The MOI is the cumulative number of infecting viruses for each host. Infected hosts lyse at rate λ. Lysis removes the infected cell and increases the local virus population by the burst size, β, regardless of the host MOI. Thus, MI effectively reduces the burst size from a single virus perspective. Additionally, we assume MI does not affect the rate of lysis (e.g., no delayed lysis).

A.3.1 Viral Dispersal

Our analysis involves varying parameters that lead to emergent host clustering in the spatial dynamics. The distance a virus diffuses before infection determines the length scale of dispersal and, in turn, the amount of host clustering occurs. We justify our choices of control parameters based on the theoretical form of the dispersal kernel that follows from classical results [104].

The dispersal kernel of an organism is the distribution of distances from which offspring occur. The dispersal of viruses is a multi-step process involving the virus diffusing to find a host followed by the infected host diffusing until lysis. We term these distributions as the adsorption kernel and lysis kernel, respectively. Each individual subprocess leads to a distribution. The total dispersal kernel is a convolution of these distributions. The adsorption kernel results from a diffusive process conditioned by viral survival and subsequent infection. Both viral decay and host infection follow exponential distribution when assuming a uniform spatial distribution of hosts. These exponential distributions convolute the

Gaussian distribution associated with diffusion into a Laplace distribution:

$$\begin{aligned} P_{\text{ads}}(x) &= \int_0^\infty (me^{-mt})(\phi N e^{-\phi N t}) \frac{e^{\frac{-x^2}{4D_v t}}}{\sqrt{4D_v t}} dt \\ &= \frac{\phi N m}{\sqrt{4D_v(\phi N + m)}} e^{-2|x|\sqrt{\frac{\phi N + m}{4D_v}}} \end{aligned} \tag{A.13}$$

where x is the location of the host after time t when starting initially at the origin. Similarly, the lysis kernel involves the host diffusing in a density dependent manner before lysing at a time following an exponential distribution. Again, the lysis kernel is conditioned on the host not dying before lysis. Hence, the lysis kernel is also a Laplace distribution:

$$\begin{aligned} P_{\text{lys}}(x) &= \int_0^\infty (de^{-dt})(\lambda e^{-\lambda t}) \frac{e^{\frac{-x^2}{4D_H(1-N/K)t}}}{\sqrt{4D_H(1-N/K)t}} dt \\ &= \frac{\lambda d}{\sqrt{4D_H(1-H/K)(\lambda + d)}} e^{-2|x|\sqrt{\frac{\lambda + d}{4D_H(1-N/K)}}} \end{aligned} \tag{A.14}$$

where $N = H + \sum_j I_j$. The overall dispersal kernel is a convolution of these two distribution yielding:

$$P_{\text{total}}(x) = \frac{b_1 e^{\frac{-|x|}{b_1}} - b_2 e^{\frac{-|x|}{b_2}}}{2(b_1^2 - b_2^2)} \tag{A.15}$$

where $b_1 = \sqrt{\frac{D_v}{\phi N + m}}$ and $b_2 = \sqrt{\frac{D_H(1-N/K)}{\lambda + d}}$. However, the adsorption kernel dominates the contribution to the total kernel for our choice of parameters since $Var(P_{lys}) \ll Var(P_{ads})$. Hence, we focus on varying parameters related to the adsorption kernel alone.

Table A.1: Parameters for spatial and mean field models. The parameters are listed in the order at which the processes occur during each time step for the spatial model. In our analysis we individually vary the adsorption rate or the viral diffusion constant while leaving other parameters constant.

Process	Symbol	Rate
host diffusion	D_H	4E-6 cm^2 hr^{-1} [6]
viral diffusion	D_v	2.04E-4 cm^2 hr^{-1} [6]
host death	d	1/(8*24) hr^{-1} [6]
viral decay	m	1/24 hr^{-1} [6]
host growth	r	1/24 hr^{-1} [6]
lysis	λ	1/12 hr^{-1} [6]
burst size	β	20 viruses [125]
adsorption	ϕ	$10^{-8.4}$ ml hr^{-1} [6]
time step	τ	1/40 hr
lattice point length	Δx	10^{-2} cm
lattice size	K	500 x 500

APPENDIX B

SUPPLEMENTARY MATERIALS FOR CHAPTER 3

B.0.1 Parameter derivations

Adsorption rates

We derive here the rate of viral adsorption. We assume that absorption and adsorption of the viral particles are diffusion limited and we solve for the rates following [126]. By solving an analogous problem of the capacitance of a dielectric sphere coated with conducting disk "receptor sites" they arrived at the following formulas for a spherical cell's intake of spherical ligands:

$$J = \ 2\pi Dac_\infty$$

where J is the maximum rate of absorption of ligands with diffusion constant D and far-off concentration c_∞. The absorbing spherical cell has diameter a. Our models use the adsorption rate, ϕ calculated from J:

$$\phi = \frac{J}{c_\infty} = 2\pi Da,$$

We use the maximum adsorption rate since we sample above and below our reference point. To estimate D we use the Stokes-Einstein Relation: $D = \frac{k_b T}{3\pi\eta d}$. This relation is relevant for spherical particles in low reynolds number fluids which is typical at micro-organismal length scales. We obtain:

$$\phi = \frac{2ak_bT}{3\eta d},$$

for a spherical molecule with diameter d in a fluid with viscosity η at temperature T where k_b is the Boltzmann constant. The relevant reference parameters in the units we used for our model are given in Table B.1.

Parameter	Meaning	Value	units	Reference
T	Temperature	293	K	-
k_b	Boltzmann cons.	1.0306×10^{-9}	$\frac{cm^2 kg}{K*day^2}$	-
η	Seawater Dynamic Viscosity	.93312	$\frac{kg}{cm*day}$	[127]
a	host diameters	$15 * 10^{-4}$ (amoeba) $3 * 10^{-4}$ (cafeteria)	cm	[128]
d_v	virus diameters	$7.5 * 10^{-5}$(mimi) $3 * 10^{-5}$(CroV)	cm	[129, 113]
d_p	virophage diameters	$7.4 * 10^{-6}$ (Sputnik) $6 * 10^{-6}$ (Mavirus)	cm	[81, 28]

Table B.1: Biophysical parameters for determining adsorption/absorption coefficients.

Host death rate, d

We chose our death rate so that the host population will grow to the carrying capacity, K, when the viral particles are absent. The dynamics in this case are:

$$\dot{H} = H(b - d(1 + \frac{H}{K})),$$

The steady-state host population is:

$$H^* = K(\frac{b}{d} - 1)$$

Hence, we choose a death rate half the value of the birth rate, $d = \frac{b}{2}$.

Virophage burst size, β_p in PEM

This value was suggested (by M. Fischer) by counting particle ratios on the electron micrographs in [78] and [76].

B.0.2 Reduced model of viral/host abundance

Here we show that including the virophage in either the IEM or the PEM effectively reduces the burst size of the virus. For the IEM we define:

$$\begin{aligned} H_{\text{total}} &= H + H_p \\ V_{\text{total}} &= V \\ \bar{\beta}_v &= \frac{\beta_v H + \beta_{vp} H_p}{H + H_p} \end{aligned}$$

Then we have that

$$\begin{aligned} \dot{H}_{\text{total}} &= \dot{H} + \dot{H}_p \\ &= H\left[b - d\left(1 + \frac{H + H_p}{K}\right)\right] + (1-\rho)bH_p - (\phi_p P + \phi_v V)H + \\ &\quad H_p\left[\rho b - d\left(1 + \frac{H + H_p}{K}\right)\right] + \phi_p H P - \phi_v V H_p \\ &= H_{\text{total}}\left[b - d\left(1 + \frac{H_{\text{total}}}{K}\right)\right] - \phi_v V_{\text{total}} H_{\text{total}} \\ \dot{V}_{\text{total}} &= \dot{V} = (\beta_v H + \beta_{vp} H_p)\frac{H_{\text{total}}}{H_{\text{total}}}\phi_v V - m_v V \\ &= \bar{\beta}_v \phi_v H_{\text{total}} V_{\text{total}} - m_v V_{\text{total}}, \end{aligned}$$

Thus for the IEM the effective dynamics for the lumped host and viral populations can be thought of as a predator-prey equations with a density dependent viral burst size, $\bar{\beta}_v$. We assume a negative effect of virophage on viral burst size such that $\beta_{vp} < \beta_v$, which gives

$$\bar{\beta}_v = \frac{\beta_v H + \beta_{vp} H_p}{H + H_p} < \frac{\beta_v H + \beta_v H_p}{H + H_p} = \beta_v$$

For the PEM define:

$$
\begin{aligned}
H_{\text{total}} &= H \\
V_{\text{total}} &= V + V_p \\
\bar{\beta}_v &= \frac{\beta_v V + (\beta_{vp} + \beta_i) V_p}{V + V_p}
\end{aligned}
$$

where the infected burst size contribution includes two parameters because viruses are part of the composite relevant to β_i. Again, we reduce the dynamics:

$$
\begin{aligned}
\dot{H}_{\text{total}} &= \dot{H} \\
&= H\left[b - d\left(1 + \frac{H}{K}\right)\right] - \phi_v (V + V_p) H \\
&= H_{\text{total}}\left[b - d\left(1 + \frac{H_{\text{total}}}{K}\right)\right] - \phi_v V_{\text{total}} H_{\text{total}} \\
\dot{V}_{\text{total}} &= \dot{V} + \dot{V}_p \\
&= (\beta_v V + \beta_{vp} V_p)\phi_v H - \phi_{vp} V P + m_p V_p - m_v V + \phi_v p V P - (m_p + m_v) V_p + \beta_i \phi_v V_p H \\
&= (\beta_v V + (\beta_{vp} + \beta_i) V_p) \frac{V_{\text{total}}}{V_{\text{total}}} \phi_v H - m_v (V + V_p) \\
&= \bar{\beta}_v \phi_v H_{\text{total}} V_{\text{total}} - m_v V_{\text{total}}
\end{aligned}
$$

Thus for both models, the dynamics for the total populations of hosts and viruses reduce to typical predator-prey dynamics with different dynamic burst sizes. An equilibrium solution to this predator prey dynamical system reveals the virophage effect on host and virus populations in the full form model:

$$
\begin{aligned}
H^*_{\text{total}} &= \frac{m_v}{\phi_v \bar{\beta}^*_v} \\
V^*_{\text{total}} &= \frac{1}{\phi_v}\left[b - d\left(1 + \frac{m_v}{\phi_v \bar{\beta}^*_v K}\right)\right]
\end{aligned}
$$

where $\bar{\beta}_v^*$ is evaluated with the respective equilibrium populations determined from the full models. Since both the IEM and PEM were reduced to the same form, this equilibrium condition holds for both models. The respective boundary equilibrium follows the same form, but with different burst size parameters:

$$\begin{aligned} H_b^* &= \frac{m_v}{\phi_v \beta_v} \\ V_b^* &= \frac{1}{\phi_v}\left(b - d\left(1 + \frac{m_v}{\phi_v \beta_v K}\right)\right) \end{aligned}$$

Since $\beta_{vp} \leq \beta_v$ and $\beta_{vp} + \beta_i \leq \beta_v$, we have $H^*_{\text{total}} \geq H_b^*$ and $V^*_{\text{total}} \leq V_b^*$. This explains why all of our stable coexistence points had higher host densities and lower virus densities when compared to their respective boundary point.

B.0.3 Parameter values for figures

Tables B.2 and B.2 present the parameters used for model simulations in Figure 3.2 and Figure B.1, corresponding to the PEM and IEM models respectively.

B.0.4 Phase lag of viruses and virophage differentiate the two models during cyclical dynamics

The order of virus and virophage peaks during cyclical dynamics may be a means for distinguishing the modes of coinfection from population level data. Specifically, the direction of the cycles in the V-P phase subspace have been found to be opposite in their orientation. We use the cyclical dynamics from Figures 3.2b,d as an example. The corresponding phase space dynamics for the IEM projected to the V-P subspace are shown in figure B.1a. The counterclockwise movement of this phase trajectory corresponds to the virophage population peak lagging the virus population peak. The phase space dynamics for the PEM projected onto the V-P subspace are shown in figure B.1b. The counterclock-

Table B.2: PEM figure parameter sets shown to 3 significant figures.

Parameter	Fig. 3.2b	Fig. 3.2d/B.1b	Fig. 3.6a	Fig. B.1c,d
b	1.84	1.53	1.84	1.50
d	0.626	0.723	0.626	0.679
K	$4.32 * 10^6$	$3.00 * 10^6$	$4.32 * 10^6$	$2.10 * 10^6$
ϕ_{vp}	$1.15 * 10^{-5}$	$4.64 * 10^{-7}$	$1.15 * 10^{-5}$	$1.64 * 10^{-6}$
ϕ_v	$3.79 * 10^{-6}$	$3.76 * 10^{-6}$	$3.79 * 10^{-6}$	$1.98 * 10^{-5}$
β_v	308	134	308	245
m_v	0.0269	0.0979	0.0270	0.146
m_p	0.297	0.0784	0.297	0.0416
ρ_p	10.5	2.20	10.5	1.59
ρ_{vp}	0.0808	0.588	0.0808	0.386
ρ_i	0.151	0.0778	0.151	0.443

Parameter	Fig. 3.2a	Fig. 3.2c/B.1a
b	1.99	1.15
d	0.862	0.913
K	$5.51 * 10^6$	$7.69 * 10^6$
ϕ_p	$5.51 * 10^{-6}$	$1.53 * 10^{-6}$
ϕ_v	$1.77 * 10^{-6}$	$8.81 * 10^{-6}$
β_v	157	162
m_v	0.0646	0.567
m_p	0.274	0.145
ρ_p	13.8	1.12
ρ_{vp}	0.343	0.348
ρ	0.392	.860

wise movement of this phase trajectory corresponds to the virus population peak lagging the virophage population peak. The same analysis was performed on a number of other points with similar results, however due to small number of examples explored, this preliminary result warrants further analysis. For example, we identified a cycle in the PEM with different dynamics when projected onto the V-P subspace as shown in Figures B.1c,d.

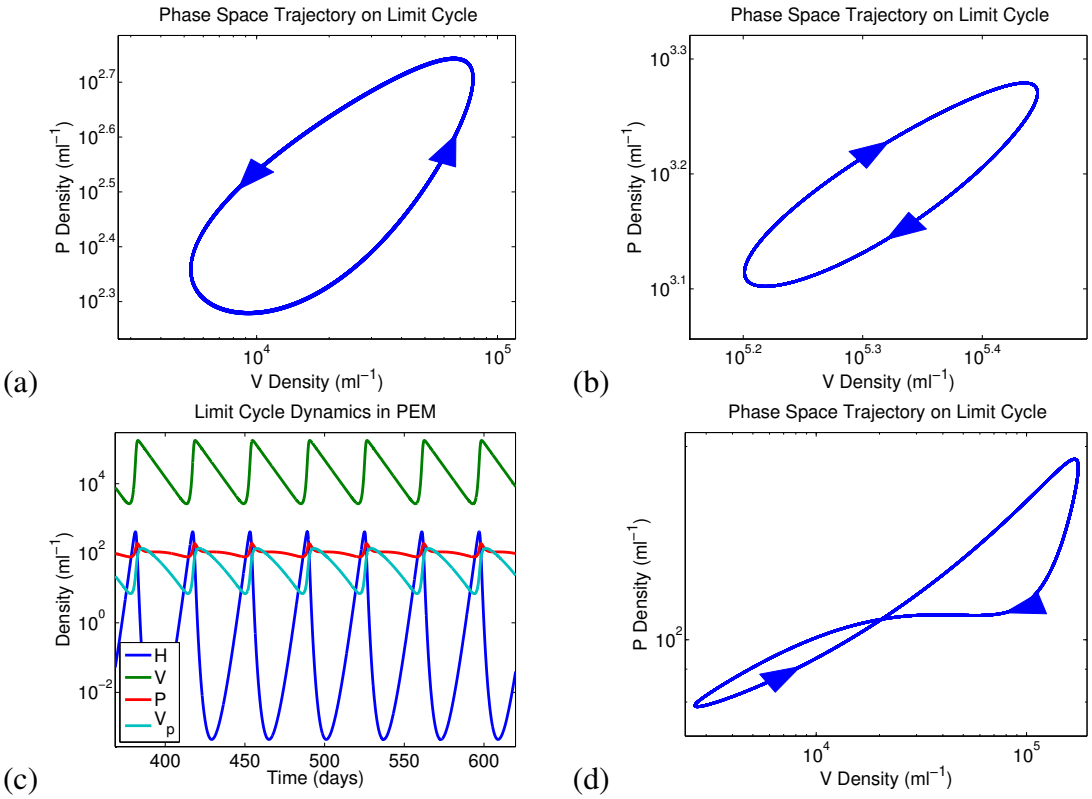


Figure B.1: Phase space representation of presented cyclical dynamics in Figures 3.2b,d for (a) IEM and (b) PEM. (c) Cyclical coexistence dynamics in PEM and respective (d) phase space representation. The parameter values are shown in Table B.2.

REFERENCES

[1] Ø. Bergh, K. Y. Børsheim, G. Bratbak, and M. Heldal, "High abundance of viruses found in aquatic environments," Nature, vol. 340, no. 6233, pp. 467–468, 1989.

[2] J. A. Fuhrman and L. Campbell, "Marine ecology: Microbial microdiversity," Nature, vol. 393, no. 6684, pp. 410–411, 1998.

[3] K. E. Wommack and R. R. Colwell, "Virioplankton: Viruses in aquatic ecosystems," Microbiology and molecular biology reviews, vol. 64, no. 1, pp. 69–114, 2000.

[4] C. A. Suttle, "Viruses in the sea," Nature, vol. 437, pp. 356–361, 2005.

[5] C. P. D. Brussaard, S. W. Wilhelm, F. Thingstad, M. G. Weinbauer, G. Bratbak, M. Heldal, S. A. Kimmance, M. Middelboe, K. Nagasaki, J. H. Paul, D. C. Schroeder, C. A. Suttle, D. Vaque, and K. E. Wommack, "Global-scale processes with a nanoscale drive: The role of marine viruses," ISME Journal, vol. 2, no. 6, p. 575, 2008.

[6] J. S. Weitz, Quantitative Viral Ecology: Dynamics of Viruses and Their Microbial Hosts. Princeton University Press, 2016.

[7] J. R. Brum and M. B. Sullivan, "Rising to the challenge: Accelerated pace of discovery transforms marine virology," Nat Rev Microbio, vol. 13, no. 3, pp. 147–159, 2015.

[8] P. G. Falkowski, T. Fenchel, and E. F. Delong, "The microbial engines that drive earth's biogeochemical cycles," Science, vol. 320, no. 5879, pp. 1034–1039, 2008.

[9] S Nee and J. M. Smith, "The evolutionary biology of molecular parasites," Parasitology, vol. 100, no. S1, S5–S18, 1990.

[10] A. H. Doermann, "Lysis and lysis inhibition with Escherichia coli bacteriophage," Journal of Bacteriology, vol. 55, no. 2, p. 257, 1948.

[11] R. Gadagkar and K. P. Gopinathan, "Bacteriophage burst size during multiple infections," Journal of Biosciences, vol. 2, no. 3, pp. 253–259, 1980.

[12] D. Refardt, "Within-host competition determines reproductive success of temperate bacteriophages," ISME journal, vol. 5, no. 9, pp. 1451–1460, 2011.

[13] P. Kourilsky, "Lysogenization by bacteriophage lambda," Molecular and General Genetics, vol. 122, no. 2, pp. 183–195, 1973.

[14] S. Luria, "Reactivation of irradiated bacteriophage by transfer of self-reproducing units," PNAS, vol. 33, no. 9, pp. 253–264, 1947.

[15] M. Meselson and J. J. Weigle, "Chromosome breakage accompanying genetic recombination in bacteriophage," PNAS, vol. 47, no. 6, pp. 857–868, 1961.

[16] A. D. Hershey and R. Rotman, "Genetic recombination between host-range and plaque-type mutants of bacteriophage in single bacterial cells," Genetics, vol. 34, no. 1, p. 44, 1949.

[17] J. S. Weitz, T. Poisot, J. R. Meyer, C. O. Flores, S. Valverde, M. B. Sullivan, and M. E. Hochberg, "Phage–bacteria infection networks," Trends in microbiology, vol. 21, no. 2, pp. 82–91, 2013.

[18] G. Hardin, "The tragedy of the commons," Science, vol. 162, no. 3859, pp. 1243–1248, 1968.

[19] M Delbrück, "Interference between bacterial viruses: Iii. the mutual exclusion effect and the depressor effect1, 2," Journal of bacteriology, vol. 50, no. 2, p. 151, 1945.

[20] P. E. Turner and L. Chao, "Prisoner's dilemma in an RNA virus," Nature, vol. 398, no. 6726, pp. 441–443, 1999.

[21] S. B. Joseph, K. A. Hanley, L. Chao, and C. L. Burch, "Coinfection rates in φ6 bacteriophage are enhanced by virus-induced changes in host cells," Evolutionary Applications, vol. 2, no. 1, pp. 24–31, 2009.

[22] M. Ptashne, "A genetic switch: Gene control and phage. lambda," 1986.

[23] T. Greiner, F. Frohns, M. Kang, J. L. Van Etten, A. Käsmann, A. Moroni, B. Hertel, and G. Thiel, "Chlorella viruses prevent multiple infections by depolarizing the host membrane," Journal of General Virology, vol. 90, no. 8, pp. 2033–2039, 2009.

[24] B. La Scola, S. Audic, C. Robert, L. Jungang, X. de Lamballerie, M. Drancourt, R. Birtles, J.-M. Claverie, and D. Raoult, "A giant virus in amoebae," Science, vol. 299, no. 5615, pp. 2033–2033, 2003.

[25] N. Philippe, M. Legendre, G. Doutre, Y. Couté, O. Poirot, M. Lescot, D. Arslan, V. Seltzer, L. Bertaux, C. Bruley, J. Garin, C. Jean-Michel, and C. Abergel, "Pandoraviruses: Amoeba viruses with genomes up to 2.5 mb reaching that of parasitic eukaryotes," Science, vol. 341, no. 6143, pp. 281–286, 2013.

[26] R. R. Novoa, G. Calderita, R. Arranz, J. Fontana, H. Granzow, and C. Risco, "Virus factories: Associations of cell organelles for viral replication and morphogenesis," Biology of the Cell, vol. 97, no. 2, pp. 147–172, 2005.

[27] Y. Mutsafi, N. Zauberman, I. Sabanay, and A. Minsky, "Vaccinia-like cytoplasmic replication of the giant mimivirus," PNAS, vol. 107, no. 13, pp. 5978–5982, 2010.

[28] M. G. Fischer and C. A. Suttle, "A virophage at the origin of large DNA transposons," Science, vol. 332, no. 6026, pp. 231–234, 2011.

[29] P. Colson, X. De Lamballerie, G. Fournous, and D. Raoult, "Reclassification of giant viruses composing a fourth domain of life in the new order megavirales," Intervirology, vol. 55, no. 5, pp. 321–332, 2012.

[30] N. Yutin, Y. I. Wolf, and E. V. Koonin, "Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life," Virology, vol. 466, pp. 38–52, 2014.

[31] B. La Scola, C. Desnues, I. Pagnier, C. Robert, L. Barrassi, G. Fournous, M. Merchat, M. Suzan-Monti, P. Forterre, E. Koonin, and D. Raoult, "The virophage as a unique parasite of the giant mimivirus," Nature, vol. 455, no. 7209, pp. 100–104, 2008.

[32] P. von Magnus, "Incomplete forms of influenza virus," Advances in virus research, vol. 2, pp. 59–79, 1954.

[33] C.-C. Hu, Y.-H. Hsu, and N.-S. Lin, "Satellite RNAs and satellite viruses of plants," Viruses, vol. 1, no. 3, pp. 1325–1350, 2009.

[34] A. M. Turing, "The chemical basis of morphogenesis," Philos Trans R Soc Lond B, vol. 237, no. 641, pp. 37–72, 1952.

[35] S. A. Levin and L. A. Segel, "Hypothesis for origin of planktonic patchiness," Nature, vol. 259, no. 5545, p. 659, 1976.

[36] R. Durrett and S. Levin, "The importance of being discrete (and spatial)," Theor. Pop. Bio., vol. 46, no. 3, pp. 363–394, 1994.

[37] D. J. Sullivan and W. Völkl, "Hyperparasitism: Multitrophic ecology and behavior," Annual review of entomology, vol. 44, no. 1, pp. 291–315, 1999.

[38] S. Yau, F. M. Lauro, M. Z. DeMaere, M. V. Brown, T. Thomas, M. J. Raftery, C. Andrews-Pfannkoch, M. Lewis, J. M. Hoffman, J. A. Gibson, and R. Cavicchioli, "Virophage control of antarctic algal host-virus dynamics," PNAS, vol. 108, no. 15, pp. 6163–6168, 2011.

[39] J. Zhou, W. Zhang, S. Yan, J. Xiao, Y. Zhang, B. Li, Y. Pan, and Y. Wang, "Diversity of virophages in metagenomic data sets," Journal of Virology, vol. 87, no. 8, pp. 4225–4236, 2013.

[40] R. K. Campos, P. V. Boratto, F. L. Assis, E. R. Aguiar, L. C. Silva, J. D. Albarnaz, F. P. Dornas, G. S. Trindade, P. P. Ferreira, J. T. Marques, C. Robert, D. Raoult, E. G. Kroon, B. La Scola, and J. S. Abrahão, "Samba virus: A novel mimivirus from a giant rain forest, the Brazilian Amazon," Virology Journal, vol. 11, no. 1, pp. 1–11, 2014.

[41] S. Oh, D. Yoo, and W.-T. Liu, "Metagenomics reveals a novel virophage population in a tibetan mountain lake," Microbes and Environments, no. 0, 2016.

[42] S. Heilmann, K. Sneppen, and S. Krishna, "Sustainability of virulence in a phage-bacterial ecosystem," Journal of Virology, vol. 84, no. 6, pp. 3016–3022, 2010.

[43] J. A. Fuhrman, "Marine viruses and their biogeochemical and ecological effects," Nature, vol. 399, no. 6736, pp. 541–548, 1999.

[44] S. Roux, A. K. Hawley, M. T. Beltran, M. Scofield, P. Schwientek, R. Stepanauskas, T. Woyke, S. J. Hallam, and M. B. Sullivan, "Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell-and metagenomics," ELife, vol. 3, e03125, 2014.

[45] S. Roux, S. J. Hallam, T. Woyke, and M. B. Sullivan, "Viral dark matter and virus–host interactions resolved from publicly available microbial genomes," ELife, vol. 4, e08490, 2015.

[46] M. A. Nowak, "What is a quasispecies?" Trends in Ecology & Evolution, vol. 7, no. 4, pp. 118–121, 1992.

[47] M. Avlund, I. B. Dodd, S. Semsey, K. Sneppen, and S. Krishna, "Why do phage play dice?" Journal of Virology, vol. 83, no. 22, pp. 11 416–11 420, 2009.

[48] A. Grande-Pérez, E. Lázaro, P. Lowenstein, E. Domingo, and S. C. Manrubia, "Suppression of viral infectivity through lethal defection," PNAS, vol. 102, no. 12, pp. 4448–4452, 2005.

[49] J. Iranzo and S. C. Manrubia, "Stochastic extinction of viral infectivity through the action of defectors," Europhysics Letters, vol. 85, no. 1, p. 18 001, 2008.

[50] E. W. Six and C. A. C. Klug, "Bacteriophage P4: A satellite virus depending on a helper such as prophage P2," Virology, vol. 51, no. 2, pp. 327–344, 1973.

[51] S. Alizon, J. C. de Roode, and Y. Michalakis, "Multiple infections and the evolution of virulence," Ecology Letters, vol. 16, no. 4, pp. 556–567, 2013.

[52] D. Phan and D. Wodarz, "Modeling multiple infection of cells by viruses: Challenges and insights," Mathematical Biosciences, vol. 264, pp. 21–28, 2015.

[53] K. W. Cummings, D. N. Levy, and D. Wodarz, "Increased burst size in multiply infected cells can alter basic virus dynamics," Biology Direct, vol. 7, p. 16, 2012.

[54] A. Asatryan, D. Wodarz, and N. L. Komarova, "New virus dynamics in the presence of multiple infection," Journal of Theoretical Biology, vol. 377, pp. 98–109, 2015.

[55] J. Yin and J. S. McCaskill, "Replication of viruses in a growing plaque: A reaction-diffusion model.," Biophysical Journal, vol. 61, no. 6, p. 1540, 1992.

[56] S. T. Abedon and R. R. Culler, "Bacteriophage evolution given spatial constraint," Journal of Theoretical Biology, vol. 248, no. 1, pp. 111–119, 2007.

[57] D. A. Jones, H. L. Smith, H. R. Thieme, and G. Röst, "On spread of phage infection of bacteria in a petri dish," SIAM Journal on Applied Mathematics, vol. 72, no. 2, pp. 670–688, 2012.

[58] V. de Rioja, J Fort, and N Isern, "Front propagation speeds of T7 virus mutants," Journal of Theoretical Biology, vol. 385, pp. 112–118, 2015.

[59] J. Aguirre and S. C. Manrubia, "Effects of spatial competition on the diversity of a quasispecies," Physical Review Letters, vol. 100, no. 3, p. 038 106, 2008.

[60] J. A. Capitán, J. A. Cuesta, S. C. Manrubia, and J. Aguirre, "Severe hindrance of viral infection propagation in spatially extended hosts," PloS One, vol. 6, no. 8, e23358, 2011.

[61] S. Heilmann, K. Sneppen, and S. Krishna, "Coexistence of phage and bacteria on the boundary of self-organized refuges," PNAS, vol. 109, no. 31, pp. 12 828–12 833, 2012.

[62] J. O. Haerter and K. Sneppen, "Spatial structure and Lamarckian adaptation explain extreme genetic diversity at CRISPR locus," MBio, vol. 3, no. 4, e00126–12, 2012.

[63] P. Roychoudhury, N. Shrestha, V. R. Wiss, and S. M. Krone, "Fitness benefits of low infectivity in a spatially structured population of bacteriophages," Proc Royal Society B, vol. 281, no. 1774, p. 20 132 563, 2014.

[64] T. Butler and N. Goldenfeld, "Robust ecological pattern formation induced by demographic noise," Physical Review E, vol. 80, no. 3, p. 030 902, 2009.

[65] M. Raghib, N. A. Hill, and U. Dieckmann, "A multiscale maximum entropy moment closure for locally regulated space–time point process models of population dynamics," Journal of Mathematical Biology, vol. 62, no. 5, pp. 605–653, 2011.

[66] J. A. Cuesta, J. Aguirre, J. A. Capitán, and S. C. Manrubia, "Struggle for space: Viral extinction through competition for cells," Physical Review Letters, vol. 106, no. 2, p. 028 104, 2011.

[67] M.-J. Lu and U. Henning, "Superinfection exclusion by T-even-type coliphages," Trends in microbiology, vol. 2, no. 4, pp. 137–139, 1994.

[68] J. Iranzo and S. C. Manrubia, "Evolutionary dynamics of genome segmentation in multipartite viruses," Proc Royal Society B, vol. 279, no. 1743, pp. 3812–3819, 2012.

[69] R. M. Donlan, "Biofilms: Microbial life on surfaces," Emerging Infectious Diseases, vol. 8, no. 9, 2002.

[70] J. R. Seymour, L. Seuront, M. J. Doubell, and J. G. Mitchell, "Mesoscale and microscale spatial variability of bacteria and viruses during a Phaeocystis globosa bloom in the Eastern English Channel," Estuarine, Coastal and Shelf Science, vol. 80, no. 4, pp. 589–597, 2008.

[71] N. Blackburn, T. Fenchel, and J. Mitchell, "Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria," Science, vol. 282, no. 5397, pp. 2254–2256, 1998.

[72] J. R. Seymour, J. G. Mitchell, L. Pearson, and R. L. Waters, "Heterogeneity in bacterioplankton abundance from 4.5 millimetre resolution sampling," Aquatic Microbial Ecology, vol. 22, no. 2, pp. 143–153, 2000.

[73] L. M. Dann, J. G. Mitchell, P. G. Speck, K. Newton, T. Jeffries, and J. Paterson, "Virio-and bacterioplankton microscale distributions at the sediment-water interface," PloS One, vol. 9, no. 7, e102805, 2014.

[74] L. M. Dann, J. S. Paterson, K. Newton, R. Oliver, and J. G. Mitchell, "Distributions of virus-like particles and prokaryotes within microenvironments," PloS One, vol. 11, no. 1, 2016.

[75] J. L. Van Etten, L. C. Lane, and D. D. Dunigan, "DNA viruses: The really big ones (giruses)," Annual Review of Microbiology, vol. 64, pp. 83–99, 2010.

[76] C. Desnues and D. Raoult, "Inside the lifestyle of the virophage," Intervirology, vol. 53, no. 5, pp. 293–303, 2010.

[77] G. Cohen, L. Hoffart, B. La Scola, D. Raoult, and M. Drancourt, "Ameba-associated keratitis, france," Emerging Infectious Diseases, vol. 17, no. 7, pp. 1306–1308, 2011.

[78] M. Gaia, I. Pagnier, A. Campocasso, G. Fournous, D. Raoult, and B. La Scola, "Broad spectrum of mimiviridae virophage allows its isolation using a mimivirus reporter," PloS One, vol. 8, no. 4, 2013.

[79] H. Ogata and J.-M. Claverie, "Microbiology - how to infect a mimivirus," Science, vol. 321, no. 5894, pp. 1305–1306, 2008.

[80] C. Xiao, Y. G. Kuznetsov, S. Sun, S. L. Hafenstein, V. A. Kostyuchenko, P. R. Chipman, M. Suzan-Monti, D. Raoult, A. McPherson, and M. G. Rossmann, "Structural studies of the giant mimivirus," PLoS Biology, vol. 7, no. 4, pp. 958–966, 2009.

[81] S. Sun, B. La Scola, V. D. Bowman, C. M. Ryan, J. P. Whitelegge, D. Raoult, and M. G. Rossmann, "Structural studies of the sputnik virophage," Journal of Virology, vol. 84, no. 2, pp. 894–897, 2010.

[82] M. Boyer, S. Azza, L. Barrassi, T. Klose, A. Campocasso, I. Pagnier, G. Fournous, A. Borg, C. Robert, X. Zhang, C. Desnues, B. Henrissat, M. G. Rossmann, B. La Scola, and D. Raoult, "Mimivirus shows dramatic genome reduction after intraamoebal culture," PNAS, vol. 108, no. 25, pp. 10 296–10 301, 2011.

[83] C. Desnues and D. Raoult, "Virophages question the existence of satellites," Nat Rev Microbio, vol. 10, no. 3, p. 234, 2012.

[84] D. Wodarz, "Evolutionary dynamics of giant viruses and their virophages," Ecology and Evolution, vol. 3, no. 7, pp. 2103–2115, 2013.

[85] M. H. Cortez and J. S. Weitz, "Distinguishing between indirect and direct modes of transmission using epidemiological time series," American Naturalist, vol. 181, no. 2, E43–E52, 2013.

[86] E Beretta and Y Kuang, "Modeling and analysis of a marine bacteriophage infection with latency period," Nonlinear Analysis-Real World Applications, vol. 2, no. 1, pp. 35–74, 2001.

[87] B. Levin, F. Stewart, and L Chao, "Resource-limited growth, competition, and predation - a model and experimental studies with bacteria and bacteriophage," American Naturalist, vol. 111, no. 977, pp. 3–24, 1977.

[88] C. Desnues, B. La Scola, N. Yutin, G. Fournous, C. Robert, S. Azza, P. Jardot, S. Monteil, A. Campocasso, E. V. Koonin, and D. Raoult, "Provirophages and

transpovirons as the diverse mobilome of giant viruses," PNAS, vol. 109, no. 44, pp. 18 078–18 083, 2012.

[89] M. G. Fischer, "Genetic and ultrastructural characterization of cafeteria roenbergensis virus and its virophage mavirus," PhD thesis, University of British Columbia, 2011.

[90] T. Byers, R. Akins, B. Maynard, R. Lefken, and S. Martin, "Rapid growth of *Acanthamoeba* in defined media – induction of encystment by glucose-acetate starvation," Journal of Protozoology, vol. 27, no. 2, pp. 216–219, 1980.

[91] R. K. Campos, K. R. Andrade, P. C. Peregrino Ferreira, C. A. Bonjardim, B. La Scola, E. G. Kroon, and J. S. Abrahao, "Virucidal activity of chemical biocides against mimivirus, a putative pneumonia agent," Journal of Clinical Virology, vol. 55, no. 4, pp. 323–328, 2012.

[92] M. Mckay, R. Beckman, and W. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," Technometrics, vol. 21, no. 2, pp. 239–245, 1979.

[93] Mathematica, Version 9.0.1.0. Champaign, Illinois: Wolfram Research, Inc., 2013.

[94] MATLAB, Version 8.0.0.783 (R2012b). Natick, Massachusetts: The MathWorks Inc., 2012.

[95] T. Kirkwood and C. Bangham, "Cycles, chaos, and evolution in virus cultures - a model of defective interfering particles," PNAS, vol. 91, no. 18, pp. 8685–8689, 1994.

[96] A. Y. Morozova, C. Robin, and A. Franc, "A simple model for the dynamics of a host-parasite-hyperparasite interaction," Journal of Theoretical Biology, vol. 249, no. 2, pp. 246–253, 2007.

[97] R. Holt and M. Hochberg, "The coexistence of competing parasites. part II - Hyperparasitism and food chain dynamics," Journal of Theoretical Biology, vol. 193, no. 3, pp. 485–495, 1998.

[98] R. May and M. Hassell, "The dynamics of multiparasitoid-host interactions," American Naturalist, vol. 117, no. 3, pp. 234–261, 1981.

[99] J. Beddington and P. Hammond, "On the dynamics of host-parasite-hyperparasite interactions," Journal of Animal Ecology, vol. 46, no. 3, pp. 811–821, 1977.

[100] B. P. Taylor, M. H. Cortez, and J. S. Weitz, "The virus of my virus is my friend: Ecological effects of virophage with alternative modes of coinfection," Journal of theoretical biology, vol. 354, pp. 124–136, 2014.

[101] U. Dieckmann and R. Law, The geometry of ecological interactions. Cambridge University Press, 2000.

[102] K. A. J. White and C. A. Gilligan, "Spatial heterogeneity in three species, plant–parasite–hyperparasite, systems," Philos Trans R Soc Lond B, vol. 353, no. 1368, pp. 543–557, 1998.

[103] M. A. Lewis, P. K. Maini, and S. V. Petrovskii, "Dispersal, individual movement and spatial ecology," Lecture Notes in Mathematics (Mathematics Bioscience Series), vol. 2071, 2013.

[104] M. G. Neubert, M. Kot, and M. A. Lewis, "Dispersal and pattern formation in a discrete-time predator-prey model," Theoretical Population Biology, vol. 48, no. 1, pp. 7–43, 1995.

[105] S. A. Frank, "Natural selection. i. variable environments and uncertain returns on investment," Journal of evolutionary biology, vol. 24, no. 11, pp. 2299–2309, 2011.

[106] K. N. Kragh, J. B. Hutchison, G. Melaugh, C. Rodesney, A. E. Roberts, Y. Irie, P. Ø. Jensen, S. P. Diggle, R. J. Allen, V. Gordon, et al., "Role of multicellular aggregates in biofilm formation," MBio, vol. 7, no. 2, e00237–16, 2016.

[107] R. H. Heineman and J. J. Bull, "Testing optimality with experimental evolution: Lysis time in a bacteriophage," Evolution, vol. 61, no. 7, pp. 1695–1709, 2007.

[108] S. T. Abedon, "Selection for bacteriophage latent period length by bacterial density: A theoretical examination," Microbial Ecology, vol. 18, no. 2, pp. 79–88, 1989.

[109] E. V. Koonin and N. Yutin, "Nucleo-cytoplasmic large dna viruses (ncldv) of eukaryotes," ELS, 2012.

[110] C. M. Brown and K. D. Bidle, "Attenuation of virus production at high multiplicities of infection in Aureococcus anophagefferens," Virology, vol. 466, pp. 71–81, 2014.

[111] M. Moniruzzaman, G. R. LeCleir, C. M. Brown, C. J. Gobler, K. D. Bidle, W. H. Wilson, and S. W. Wilhelm, "Genome of brown tide virus (Aav), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host–virus coevolution," Virology, vol. 466, pp. 60–70, 2014.

[112] M. G. Fischer and T. Hackl, "Genome integration and reactivation of the virophage mavirus in the marine protozoan cafeteria roenbergensis," BioRxiv, 2016. eprint: http://biorxiv.org/content/early/2016/08/07/068312.full.pdf.

[113] M. G. Fischer, M. J. Allen, W. H. Wilson, and C. A. Suttle, "Giant virus with a remarkable complement of genes infects marine zooplankton," PNAS, vol. 107, no. 45, pp. 19 508–19 513, 2010.

[114] C. P. D. Brussaard, J. P. Payet, C. Winter, and M. G. Weinbauer, "Quantification of aquatic viruses by flow cytometry," Manual of aquatic viral ecology, vol. 11, pp. 102–107, 2010.

[115] S. Kaczanowski, M. Sajid, and S. E. Reece, "Evolution of apoptosis-like programmed cell death in unicellular protozoan parasites," Parasites & vectors, vol. 4, no. 1, p. 1, 2011.

[116] T. Fenchel, Ecology of Protozoa: The biology of free-living phagotropic protists. Springer-Verlag, 2013.

[117] J. M. Rose, D. A. Caron, M. E. Sieracki, and N. Poulton, "Counting heterotrophic nanoplanktonic protists in cultures and aquatic communities by flow cytometry," Aquatic Microbial Ecology, vol. 34, no. 3, pp. 263–277, 2004.

[118] E. L. Ellis and M. Delbrück, "The growth of bacteriophage," The Journal of general physiology, vol. 22, no. 3, pp. 365–384, 1939.

[119] W. Bode, "Lysis inhibition in Escherichia coli infected with bacteriophage T4," Journal of virology, vol. 1, no. 5, pp. 948–955, 1967.

[120] D. Anderson and R. Watson, "On the spread of a disease with gamma distributed latent and infectious periods," Biometrika, vol. 67, no. 1, pp. 191–198, 1980.

[121] S Halary, S Temmam, D Raoult, and C Desnues, "Viral metagenomics: Are we missing the giants?" Current opinion in microbiology, vol. 31, pp. 34–43, 2016.

[122] M. Krupovic and V. Cvirkaite-Krupovic, "Virophages or satellite viruses?" Nat Rev Microbio, vol. 9, no. 11, pp. 762–763, 2011.

[123] K. A. S. Thompson and J. Yin, "Population dynamics of an rna virus and its defective interfering particles in passage cultures," Virology journal, vol. 7, no. 1, p. 1, 2010.

[124] M. Breitbart, P. Salamon, B. Andresen, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam, and F. Rohwer, "Genomic analysis of uncultured marine viral communities," PNAS, vol. 99, no. 22, pp. 14 250–14 255, 2002.

[125] R. T. Noble, M. Middelboe, and J. A. Fuhrman, "Effects of viral enrichment on the mortality and growth of heterotrophic bacterioplankton," Aquatic Microbial Ecology, vol. 18, no. 1, pp. 1–13, 1999.

[126] H. Berg and E. Purcell, "Physics of chemoreception," Biophysical Journal, vol. 20, no. 2, pp. 193–219, 1977.

[127] G. W. C. Kaye and T. H. Laby, Tables of Physical & Chemical Constants, 16th. Longman, 1995.

[128] R. Massana, J. del Campo, C. Dinter, and R. Sommaruga, "Crash of a population of the marine heterotrophic flagellate cafeteria roenbergensis by viral infection," Environmental Microbiology, vol. 9, no. 11, pp. 2660–2669, 2007.

[129] C. Xiao, P. Chipman, A. Battisti, V. Bowman, P Renesto, D Raoult, and M. Rossmann, "Cryo-electron microscopy of the giant mimivirus," Journal of Molecular Biology, vol. 353, no. 3, pp. 493–496, 2005.